water

MDPI

Review

A Systematic Review of the State of Cyber-Security in Water Systems

Nilufer Tuptuk [1,*], Peter Hazell [2], Jeremy Watson [3] and Stephen Hailes [1]

[1] Department of Computer Science, University College London, London WC1E 6EA, UK; s.hailes@ucl.ac.uk
[2] Department of Technology, Enterprise, Change and Data Science—Information & Cyber-Physical Security, Yorkshire Water, Bradford BD6 2SZ, UK; peter.hazell@yorkshirewater.co.uk
[3] Department of Science, Technology, Engineering and Public Policy, University College London, London WC1E 6BT, UK; jeremy.watson@ucl.ac.uk
* Correspondence: n.tuptuk@ucl.ac.uk

Abstract: Critical infrastructure systems are evolving from isolated bespoke systems to those that use general-purpose computing hosts, IoT sensors, edge computing, wireless networks and artificial intelligence. Although this move improves sensing and control capacity and gives better integration with business requirements, it also increases the scope for attack from malicious entities that intend to conduct industrial espionage and sabotage against these systems. In this paper, we review the state of the cyber-security research that is focused on improving the security of the water supply and wastewater collection and treatment systems that form part of the critical national infrastructure. We cover the publication statistics of the research in this area, the aspects of security being addressed, and future work required to achieve better cyber-security for water systems.

Keywords: smart water systems; cyber–physical security; cyber-security; cyber–physical attacks

check for updates

Citation: Tuptuk, N.; Hazell, P.; Watson, J.; Hailes, S. A Systematic Review of the State of Cyber-Security in Water Systems. *Water* **2021**, *13*, 81. https://dx.doi.org/10.3390/w13010081

Received: 30 November 2020
Accepted: 25 December 2020
Published: 1 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Water is becoming scarcer. According to the United Nations World Water Development Report published in 2018 [1], nearly half the world's population, around 3.6 billion people, face water-scarcity for at least one month per year, and it is expected that over 5 billion people will suffer some water shortage by 2050. The World Bank estimates that around 45 million cubic meters of water are lost each day in developing countries, costing over US$3 billion per year [2]. This loss is mainly due to inefficient infrastructure, ageing infrastructure that leaks, and non-revenue water due to lack of billing or inaccuracies in costing such as metering issues [2]. It affects both developed and developing countries. In England and Wales 2954 million litres of water are leaked each day from distribution networks and supply pipes [3].

Climate change, water pollution, increasing urbanisation and population growth, ageing and inefficient infrastructure, compliance with tighter regulation and water quality standards are some of the challenges faced by water sector in seeking to maintain their services. To resolve these challenges, water and wastewater providers are moving towards smart water systems [4–6] that are reliable, efficient and that support real-time decision-making. This is particularly true in the UK, where the UK government has established strategic priorities for the period from 2020 to 2025 aimed at securing long-term resilience in the water industry; these are supported by major investments by water companies and providers [7,8].

Water systems are a type of cyber–physical system (CPS) that integrate computational and physical capabilities to control and monitor physical processes. In the past, water system security was achieved largely through isolation, limiting access to control components. However, with the emergence of IoT, water systems, as with other critical infrastructure services, are increasingly using a smart systems philosophy. This promotes

the incorporation of IoT and analytics into industrial control systems (ICS) to improve the sensing and control capacity and ensure better integration with business processes. Collectively, this is known as the Industrial Internet of Things (IIoT), often labelled Industry 4.0, in which IoT is applied to industrial applications. It relies on connecting multiple layers of cyber–physical systems to facilitate autonomous decentralised decision-making and to improve the use of real-time data and predictive analytics to promote reliability, efficiency and productivity. With these technological advances, water systems that collect, treat, transport and distribute water to customers are undergoing a similar transformation, becoming highly connected and facing new technological challenges in the drive to provide safe water reliably.

ICS deployment often follows a hierarchical architectural approach that is sometimes characterised using the Purdue reference model [9], as shown in Figure 1. This spans multiple layers, encompassing the variety of equipment and communication protocols and the range of goals and complexity that are likely to be found in these environments [9].

Level 5, the enterprise network, is the level at which business decisions are made, and in which the regular corporate systems (enterprise desktops and servers) operate. At Level 4, the site business planning and logistics applications and systems are found. At Level 3, the operations network, operations management systems such as domain controllers, data collection servers (historians) and application servers are found. Level 2, supervisory control, consists of devices that monitor and control the process at the lower levels. Typically, these consist of supervisory interfaces for the operators, engineering workstations, and distributed control servers that monitor and control various parts of production. At Level 1, controllers monitor and control a set of devices autonomously and/or based on decisions that come from the supervisory system. They receive inputs from instrumentation equipment (e.g., field devices) such as sensors, and send output signals to other devices (actuators). Level 0 is where the actual process takes place, containing the sensors and actuators connected via a fieldbus network.

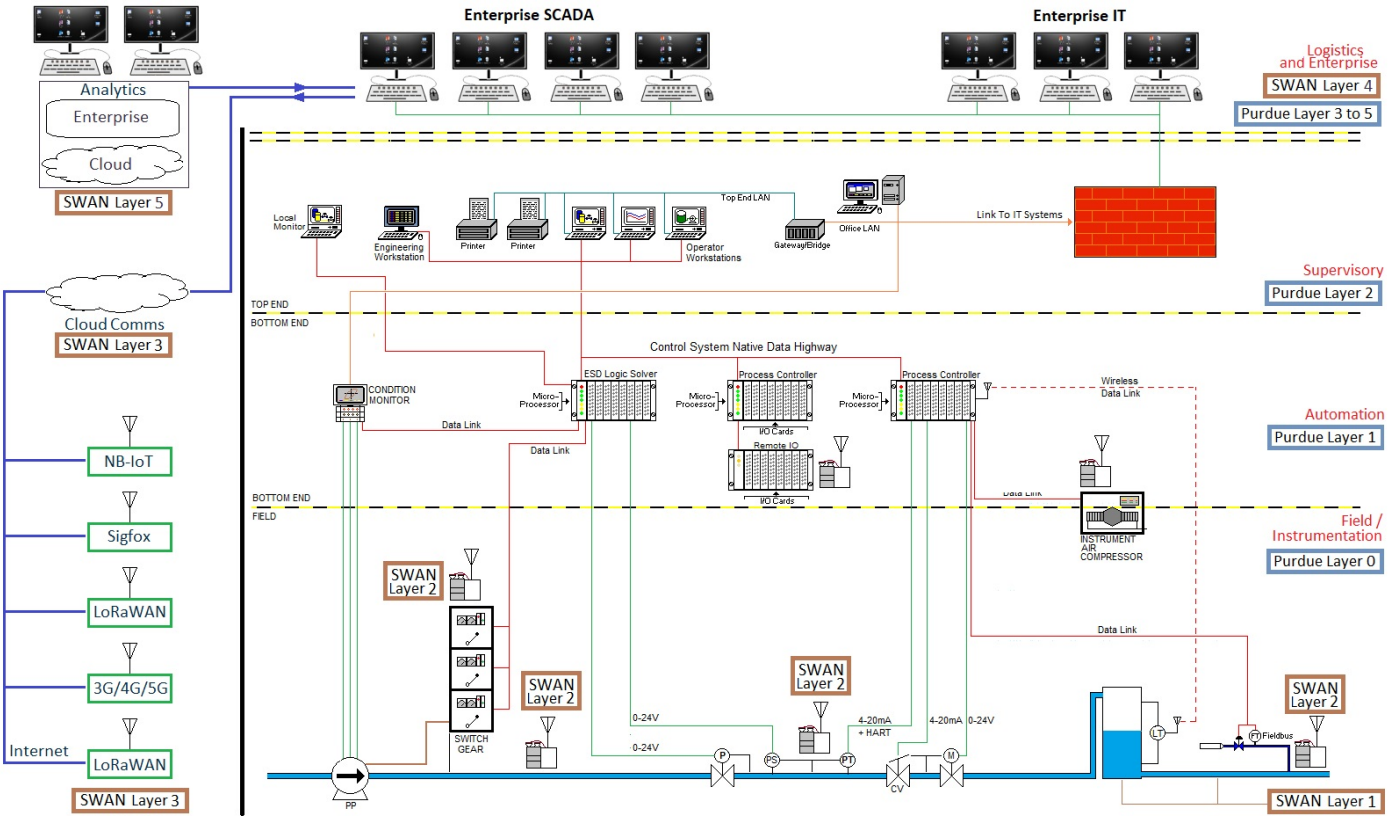

Figure 1. Purdue reference model with SWAN layers.

According to the Smart Water Networks Forum (SWAN) [10], a global non-profit hub consisting of international water companies, academics, regulators, and other water experts, smart water networks are the "entire system of data technologies connected to or serving the water distribution network [and] it is informative to separate its components into layers." These layers [10] are similar to those found in Purdue reference model, as indicated in Figure 1:

- Level 1: Physical layer is composed of physical devices that provide the distribution and delivery of water services. This includes pipes, pumps, valves, reservoirs and endpoints for delivering water.
- Level 2: Sensing and control layer is composed of equipment and sensors responsible for gathering measurements for monitoring and controlling water delivery and distribution; and remote-controlled actuators to remotely operate water networks.
- Level 3: Collection and communications layer provides the data collection, transmission, and storage between layer 2 and level 4 where the instructions for sensors and actuators are computed. All network protocols used for data transfer are found in this layer.
- Level 4: Data management and display layer is responsible for gathering and managing data from different sources. Supervisory control and data acquisition (SCADA) systems, control systems, visualisation systems and tools such as human-machine interface (HMI), data storage repositories and control systems are found in this layer. This is where decisions taken by upper layers are interpreted into control and other commands such as settings for devices at lower layers.
- Level 5: The data fusion and analysis layer is where raw data is processed into information and where the "smart" emerging technologies are deployed. These include modelling and optimisation systems, network infrastructure monitoring, and other supporting and decision support systems for managing water networks.

The adoption of network communication, the increasing use of commercial-off-the-shelf (COTS) components and the deployment of wireless systems in Purdue and SWAN architecture layers bring new security challenges as they have the potential to expose water systems to a wide variety of adversaries. The number of reported attacks targeting cyber–physical systems that are critical for national infrastructure services has been on the increase and, as the evidence from successful attacks such as Stuxnet [11], DuQu [12], BlackEnergy [13] and Havex [14] shows, such attacks can have catastrophic consequences. The criticality of water to human life and the ecosystem means that water systems are an obvious target for political, military and terrorist actors [15,16].

Table 1 reports some of the incidents against water infrastructure services that have been made public. These indicate the potential for successful attacks to exploit a wide variety of vulnerabilities and so cause both direct disruption of services and damage to control equipment and communication networks that, in turn, may affect essential services. The broader impacts of such attacks lie in the health of both the public and the ecosystem, as well as in financial and reputational losses for the companies affected. Hassanzadeh et al. [17] report a review of 15 water incidents, including some of the attacks summarised in Table 1. A widely referenced source for cyber-security incidents in the water sector is the work carried out by Industrial Control Systems Cyber Emergency Response Team (ICS-CERT) in the United States. This tells us that, in 2015, the US Department of Homeland Security (DHS) recorded 25 cyber-security incidents from the water sector [18].

Table 1. Past attacks on water systems.

Reference	Year	Target	Attribution	Infection Vector	Details	Impact
Israel's water system [19]	2020	OP	Hacktivist/ Nation state	Unknown	Israeli government reported cyber-attacks against water supply and treatment facilities and urged these facilities to change passwords.	Unknown.
Northern Colorado [20]	2019	OP	Cybercrime	Ransomware	Locked access to technical and engineering data.	Disruption, took about three weeks to unlock data.
Cryptojacking [21]	2018	OP	Cybercrime	Cryptocurrency mining	Cryptocurrency malware installed on HMI on the SCADA network.	Unknown.
Kemuri water [22]	2016	OP	Hacktivist	Remote access	Accessed PLC responsible for controlling water treatment chemicals.	Engineers were able to identify and reverse the changes made to process control parameters.
Bowman Avenue Dam [23,24]	2016	OP	Hackers/ Nation state	Remote access	According to US authorities, hackers linked to Iranian Armed Forces infiltrated ICS of Bowman Avenue Dam and accessed the SCADA for the dam.	Data exfiltration and over $30k on remediation costs. Physical damage was not possible due to disconnected sluice gates.
Florida Wastewater [25]	2012	IT	Ex-Employee	Remote access	Stolen login credentials were used to access district's computer system.	Deleting and modifying information. Ex-employee was arrested on account of computer crime.
Tehama-Colusa Canal [26]	2007	OP	Ex-employee	Physical access	Installed malware on SCADA system responsible for controlling agricultural irrigation [26].	Damage to equipment, and additional unknown amount of monetary loss due to replacing production.
Harrisburg water plant [27]	2006	IT	Hackers	Remote Access	Compromised and installed malware on an employee's laptop which could have been used as an entry point to reach water treatment system.	Unknown.
Maroochy Shire [28,29]	2000	OP	Ex-employee of a contractor	Physical access	Masqueraded as a controller using stolen equipment and sent fake commands to the pumping station.	Approximately 800,000 litres of sewage was released into the environment, harming local parks and rivers, impacting public health, killing marine life, and caused large monetary loss.

Cyber-attacks against infrastructure services are often not made public and attribution of these incidents can be a complex and uncertain process, requiring well-developed skills and capabilities [30] to identify the actors. Nevertheless, publicly reported incidents show that the sources of cyber-attacks against water systems appear to include a wide variety of actors. These include hacktivists who perform cyber-attacks often based on a political ideology; disgruntled former employees seeking revenge; cybercriminal networks motivated by monetary gain; and hacker hobbyists who attack for fun, curiosity, or the desire for recognition [31]. Other potential adversaries include nation-state-sponsored attacks for political gain and industrial espionage; rival organisations or companies seeking business advantage; terrorist groups attacking national security; and insiders motivated by problems at work, political or monetary gain, fear/coercion or just for the thrill or fun.

The current history of incidents suggests that the design and performance of advanced targeted attacks against operational processes (OP) require actors with more than just IT skills [32]. Until recently, most of the cyber-attacks against cyber–physical processes were carried out by insiders, with most of the remainder conducted by nation states. In other words, most attacks have been conducted by those with the knowledge, skills and resources needed to cause a real physical impact. More recently, however, there has been

an increasing incidence of cyber-criminals targeting industrial processes, with the aim of installing ransomware [33].

In this paper, we present a systematic literature review and evaluate the current state of cyber-security of cyber–physical systems within the water sector, focusing on process control layers, as the corporate IT layers are primarily affected by security problems covered by traditional information security. Our aim is to identify what is being done, by whom, where, how and what aspects of cyber-security are being covered.

The remainder of this paper is structured as follows. Section 2 provides brief overview of cyber–physical system security. Section 3 describes the research questions and methodology used for carrying out the systematic review. Key research findings are reported and discussed in Section 4. Section 5 highlights the limitations of existing studies and discusses some direction for future research. Finally, Section 6 concludes the paper.

2. Cyber–Physical Systems

The term "cyber–physical system" (CPS) was first coined by Helen Gill at the National Science Foundation (NSF) in 2006 to describe "physical, biological and engineered systems whose operations are integrated, monitored, and/or controlled by a computational core" [34]. Since then, CPS have attracted significant research effort, including initiatives in Industry 4.0, the Internet of Things and the Industrial Internet of Things. As computer scientist Edward A. Lee points out [35], terms such as the Internet of Things (IoT), Industry 4.0, the Industrial Internet (II), Machine to Machine (M2M), the Industrial Internet of Things (IIoT) and other similar terms have been strongly connected with CPS, and sometimes used interchangeably and sometimes for specific sectors (e.g., Industry 4.0 for manufacturing). However, these terms cover "implementation approaches (e.g., the "Internet" in IoT) or particular applications (e.g., Industry 4.0)" [35]. CPS are found in a broad range of sectors including health care and medicine, materials, manufacturing, automotive, aerospace, utilities, chemical, civil infrastructure and transportation [34]. Despite the differences in interpretation, many industry sectors share common technologies and, by extension, share similar concerns relating to their security. A common concern for all these sectors in adopting new enabling technologies for CPS is to ensure security in the face of cyber-attacks.

2.1. Securing Cyber–Physical Systems

The National Institute of Standards and Technology (NIST) defines cyber-security as "the process of protecting information by preventing, detecting and responding to attacks" [36]. The prevention of attacks against information technology systems is defined in terms of three security goals: confidentiality, integrity and availability, known as the CIA triad. These goals are also applied to CPS to maintain security.

Confidentiality ensures data or system resources "are not disclosed to unauthorised individuals, processes, or devices" [37]. The operation of CPS requires, *inter alia*, data from instrumentation devices, controllers, supervisory control systems, monitoring and safety systems. Unauthorised access to this data is potentially useful for preparing and implementing attacks and for industrial espionage. Integrity deals with "guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity" [38]. Violating integrity could interfere with the operation of CPS and undermine the reliability and safety of the CPS process. Availability deals with "timely, reliable access to data and information services for authorised entities" [39]. Many CPS are continuous systems and loss of availability can cause systems to shut down and interrupt the production process. Usually, integrity and availability are the most important concern for critical cyber–physical systems [40], but the priority given to each of these security goals depends on the risks associated with loss of these properties in the context of a particular system.

Cyber–physical systems have control properties that need to be maintained. These include stability, observability, controllability, safety and efficiency [41], as well as accuracy,

responsiveness, rapid disturbance rejection and low control effort. Security attacks aimed at sabotaging CPS involve the manipulation of these properties; thus, the maintenance of these properties, even when the system is under attack, is an essential component of ensuring the security of CPS.

2.2. Attacks against Cyber–Physical Systems

Figure 2 shows the typical components of a networked CPS. The controller is given a process reference (Setpoint-SP) as the desired process output to maintain. The sensor measures the output of the physical process (Measured Process Value-PV) and sends this over a network to the controller. The controller (for example a PLC) receives these values, compares the PV against the desired SP reference value, calculates a control command (Manipulated Variable-MV) and sends this, through the network, to the actuator. The actuator acts on this command and outputs a physical control action that modifies the process. Attacks against CPS involve attacking components of CPS to achieve either data exfiltration, which involves gathering sensitive information about the CPS, or sabotage, which involves disrupting the process.

Adversaries use a range of tools to carry out attacks against elements of Figure 2. These include attacks that compromise sensors, actuators and controllers to modify their settings or configurations so that incorrect signals are sent to relevant components; for example, incorrect control commands from controller to actuator or incorrect PVs from sensor to controller. Attacks can be carried out against the network: modifying the data in transit (replaying old data, dropping data, injecting false data); denying or delaying the flow of data (e.g., DoS, jamming attacks); or impersonating another actor (for example IP and ARP spoofing and communication hijacking). Eavesdropping attacks against networks can be carried out to gather information related to the operation of CPS, such as identifying communication protocols, open ports, hosts and applications, and sniffing network traffic. Physical attacks can be carried out against CPS components, e.g., to modify the location of devices; change device calibration; install rogue devices on the network; install malware via portable devices (e.g., USB sticks); cause changes in sensor values by manipulating the physical environment of the devices; and cause physical damage to devices.

The success of an attack depends on the resources and skills available to adversaries as well as system vulnerabilities and the absence of appropriate independent layers of protection designed to prevent mal-operation due to operator error, random equipment failure or cyber-attack. Vulnerabilities are typically introduced into CPS due to: poor security design; insecure network communication protocols; insecure backdoors and holes in the virtual or physical network perimeter; insecure software and hardware; poor management of security or ineffective policies and inappropriate physical access [40]. To exploit a CPS, a highly motivated adversary with high skills and resources can purchase zero-day vulnerabilities that are, by definition, not yet public, as seen in the past (e.g., Stuxnet [11]).

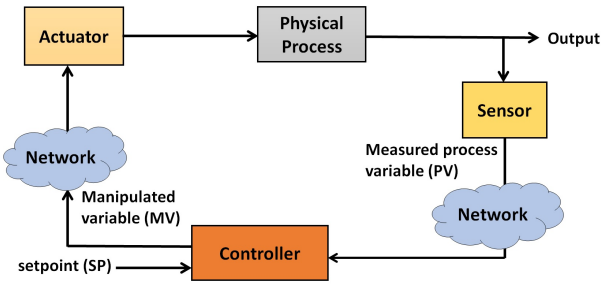

Figure 2. Typical cyber–physical system.

Adversaries have a wide variety of motivations, and impact goals depend on these motivations. Potential impacts include process disruption; damage to production, equip-

ment, safety and the environment; data disclosure; data loss; disruption to assets; injuries and loss of life; damage to reputation; and financial damage.

2.3. Security Measures for Cyber–Physical Systems

Security mechanisms to protect systems against malicious behaviour can be divided into three main categories: *preventive*, *reactive* and *responsive* measures. *Preventive* measures are security controls implemented to prevent attacks such as authentication; access control; network segmentation; maintaining confidentiality and integrity of transmitted data and in storage using cryptographic techniques; patching software vulnerabilities; deploying usable and effective security management policies that defines roles and procedures for managing and maintaining security; personnel awareness and training programs to understand threats; and measures for protecting the supply chain [40]. *Reactive* or *detection-based* measures are security controls implemented to identify attacks and anomalous behaviour such as intrusion/anomaly-based monitoring and detection for process and host; antivirus and other malware monitoring tools; and safety management systems. After an attack is detected, *response* strategies include measures to reduce damage; for example, reconfiguring the network; restricting access to network; systems or devices; deploying designed-in redundancies; and shutting down the system.

3. Methodology for Systematic Review

Our aim in this paper is to review and gain an understanding of cyber-security research targeted at protecting cyber–physical systems in the water sector, thence to identify areas that require future research. The Preferred Reporting Items for Systematic Reviews (PRISMA) [42] guidelines were followed, as illustrated in Figure 3. A set of question research questions were devised to analyse and evaluate the relevant publications. A set of electronic databases and a search strategy was designed to identify the publications. Inclusion and exclusion criteria were used to assess the eligibility of each publication. The eligible publications were then manually inspected to extract relevant evidence for analysis.

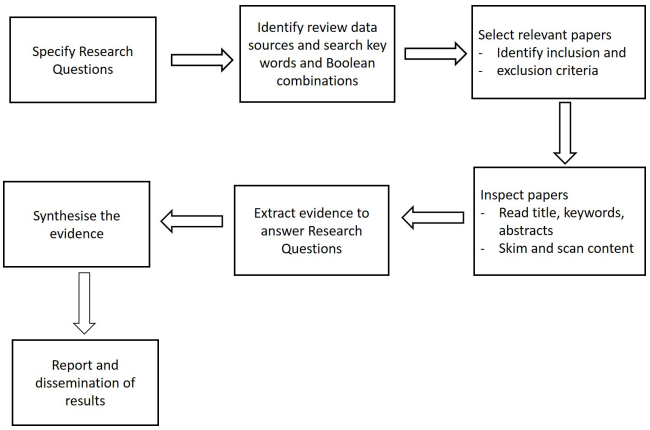

Figure 3. Systematic literature review process, adapted from [42].

3.1. Research Questions

To identify, classify and evaluate the existing cyber-security work within water sector, a set of research questions were identified.

- *RQ1 How did the number of publications change over the years?* To understand the publication trends over the years, and to understand if the topic is gaining more research focus with moves towards IIoT and Industry 4.0. Answering this question might also enable us to see any trends that might have motivated more work from the research community.
- *RQ2 What is the geographic distribution of these studies?* To understand by whom and from where these studies are being conducted. Answering RQ2 will help to determine

countries investing the least and most in research in these areas, and why this could be the case. Security of national infrastructure services such as water often require a joint effort from academia, governmental bodies and industry.

- *RQ3 What is the distribution of academic, governmental and industry studies?* To identify the level of involvement, and the support of government and industry in research studies. Answering this question will enable assessment of whether relevant government and industry bodies are participating in these studies. Their involvement is crucial for these studies, as they are essentially the clients that will deploy and implement security solutions.
- *RQ4 What are the target venues for publishing these studies?* To identify publication venues targeted by these studies. Answering this question will help to identify the top target venues for publication, and gain some understanding of the maturity and quality of publications by analysing the rating of the journals and conferences.
- *RQ5 Which security aspects are covered in these studies?* To understand the security themes of interest, proposed solutions and focus of these studies. Answering this question will inform the security problems that are being solved.
- *RQ6 Can one classify security aspects in RQ5 further?* To see if there are popular areas of research that can be classified further. If there are popular research aspects, answering this question could help to compare different approaches.

3.2. Identification of Sources and Search Term

The search strategy for identifying publications was primarily through online databases: Springer Link, IEEE Xplore, ACM, Science Direct and ASCE library. These are the most common libraries for publishing conference proceedings and journal publications within the field of cyber-security in cyber–physical systems. Google Scholar returned articles that were covered in these databases; however, we also used it to identify relevant publications that appeared in other databases or venues. The search strings used for the databases were "water and cyber-security" or "cyber-security". Table 2 shows the search string for each database. When a basic search on databases returned many papers, advanced searching was used to filter irrelevant papers. For example, searching Google Scholar using combinatorial search keywords such as "water" AND "cyber-security" resulted in a high number of papers (over 17,900) that were not relevant to this systematic review. Instead, the search was limited to terms appearing in the title: "water" and "cyber" to identify studies that primarily focused on cyber-security of water systems. A list of security keywords was also used in conjunction, to search the databases for relevant publications. These qualifiers included: water, integrity, confidentiality, availability, integrity, authentication, authorisation, access control, threat, vulnerabilities, attacks, and detection. However, these failed to capture any new publications. Searching was limited to publications that had been published from 2000 to 2020.

Table 2. Search string used for each data source.

Source	Search String
Springer	where the title contains: Water AND with at least one of the words: cyber-security OR cybersecurity
ACM Digital Library	[Document Title: water] AND [[Abstract: cyber-security] OR [Abstract: cybersecurity]]
IEEE Xplore	"All Metadata": water cyber-security
ScienceDirect	Find articles with these terms: cyber-security OR cybersecurity, title, abstract, keywords: water
ASCE Library	water AND (cyber-security OR cybersecurity)
Google Scholar	allintitle: water cyber

Figure 4 shows the number of publications retrieved from online databases. Duplicates were removed from this pool of publications and the remaining publications were included for further review.

To complement online database searching, a manual review of reference lists of eligible papers and any notable journals (e.g., Water and Environment Journal), conferences (e.g., World Environmental and Water Resources Congress) and workshops (e.g., International Workshop on Cyber–Physical Systems for Smart Water Networks) was carried out to identify any relevant publications that might have been missed in the database search.

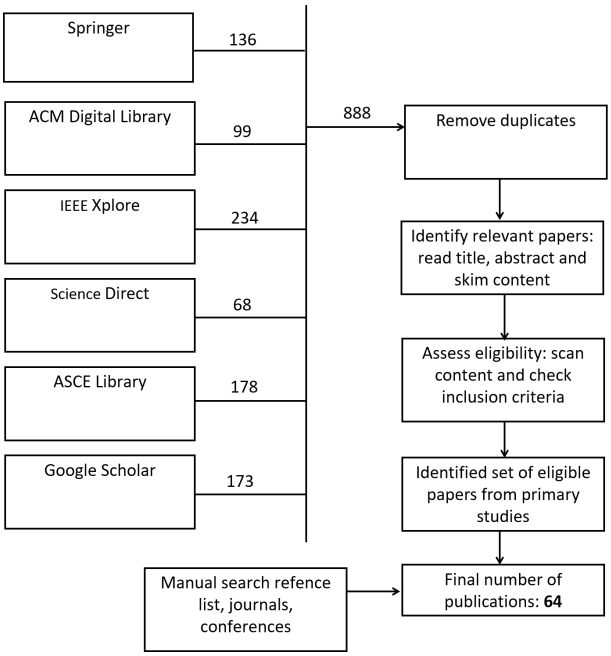

Figure 4. Publication selection process.

3.3. Criteria for Selection of Papers

Selection criteria for identifying publications for systematic review were as follows:

- Must address cyber–physical systems in water.
- Must have a technical content and address cyber-security.
- Must be peer-reviewed and must have appeared in an international journal, conference or workshop.

Books, book chapters, theses, editorials, feature or opinion pieces, essays, governmental and industry guidelines, other non-peer-reviewed or non-research publications, non-English publications, and publications appearing in local conferences, workshops or journals were excluded from the search. Review papers were not included in the analysis, but their content was analysed in the manual reference search and, where relevant, they are mentioned.

3.4. Paper Inspection

Online database searching resulted in 888 publications, and details of these were exported into a CSV file for further processing. After removing any duplicates, the remaining peer-reviewed publications published in internationally recognised conferences, workshops or journals were selected for further inspection. Selection of the eligible list of publications for analysis was based on inclusion and exclusion criteria by inspecting title and abstract, and text skimming. As a result, a set of 64 publications was finalised for analysis to answer the research questions.

3.5. Extraction of Appropriate Information

To analyse the content of the publications, the reviewed publications were classified into categories according to application domains, date of publication, number of citations, publication type, publication venue, affiliation, authors' countries of affiliation, and security aspects covered by the publication. Citation numbers for publications retrieved through online databases were not always accurate, so Google Scholar was used as a cross reference to retrieve the citation numbers. The data extracted was recorded in an Excel spreadsheet to facilitate analysis.

4. Analysis of Results

4.1. Publication Trends

Figure 5 shows the application domains of the security studies. The majority of studies were carried out on drinking water systems: 39 studies focused on security of water distribution systems (WDS) including water distribution networks; 3 studies included water supply and distribution systems; and 2 studies focused on water supply systems. Another 16 studies investigated security of drinking water treatment systems. Only four studies focused on non-drinking water systems: 3 studies focused on canal automation systems used for irrigation; and one study covered wastewater systems. There is a clear imbalance between studies covering water systems designed to provide drinking water versus those designed for other forms of water.

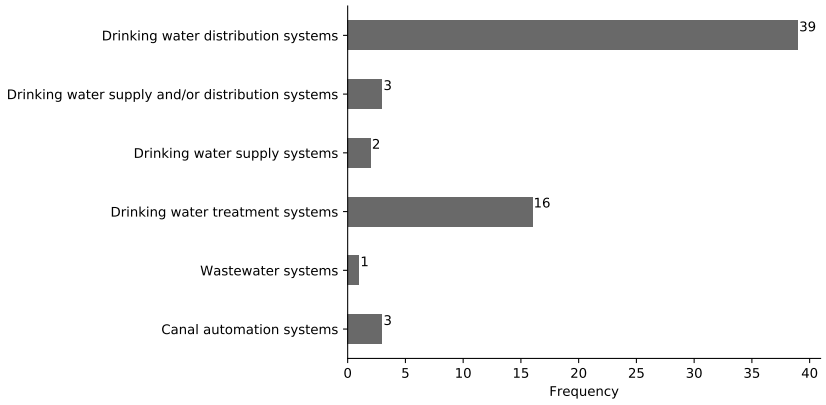

Figure 5. Application domains.

Figure 6 shows the timeline of publication. The earliest publication found dates from 2004, but most of the research effort (56 papers) was published after 2015. Answering **RQ1**, there has been increasing interest in the security of water systems over the years, likely as a result of the emergence of new resources and corresponding effort that made use of them.

These resources include the deployment of two important testbeds: the Secure Water Treatment (SWaT) testbed [43] and water distribution testbed (WADI) [44], and associated datasets [45] at the iTrust Centre for research in cyber-security at Singapore University of Technology and Design [46], and the BATADAL (BATtle of the Attack Detection Algorithms) competition organised by iTrust center and their international collaborators [47] to detect cyber-attacks against water distribution systems (WSD). This corresponds to a period (post 2016) in which associated open-source attack detection has become more available and European Commission (EC) projects such as FACIES (Online identification of Failure and Attack on interdependent Critical InfrastructurES) [48] and STOP-IT [49] have been investigating physical and cyber-security of critical water infrastructures. This trend is supported by the number of publications per country involved in these projects.

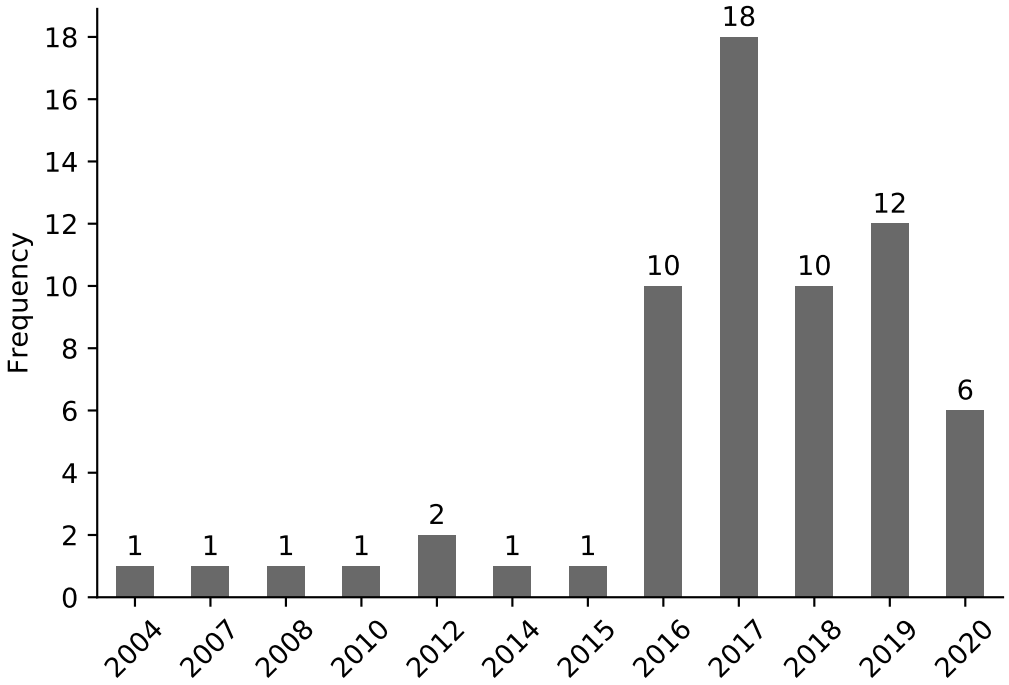

Figure 6. Number of publications over the years.

Figure 7 shows the distribution of studies per country based on the location of the authors. If the authors of the publication were located in multiple countries, for example several authors from Singapore and one author from Israel, both countries were added to the statistics. Figure 7 provides an answer to **RQ2** indicating that most of the existing research has been carried out by authors in countries that have made investments in this area: Singapore and their collaborators (Israel, USA) and countries involved in projects funded by the EC.

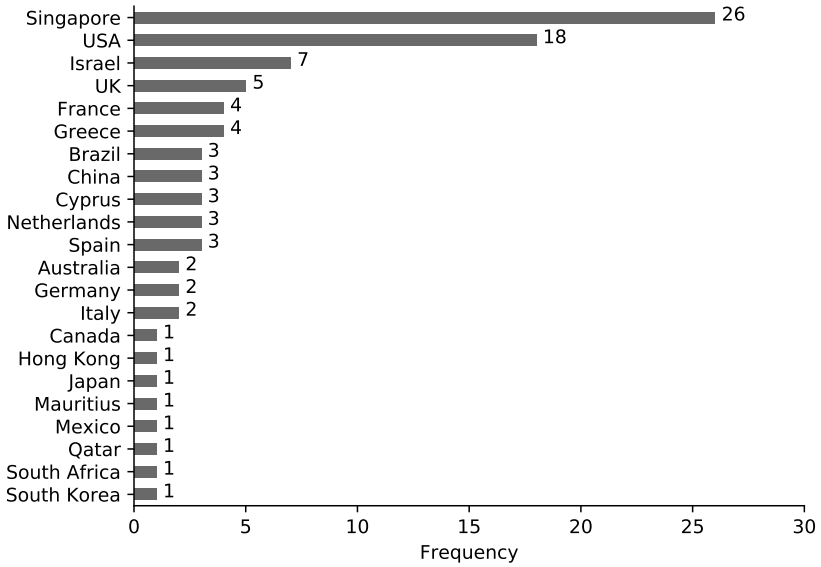

Figure 7. Country of publication based on location of authors.

Figure 8 shows the results to answer **RQ3**. Most of the research has been carried out by academia (85.1%); 6.8% was based in private organisations that provide security consulting services; 6.8% is provided by independent or public funded research organisations; and one

paper (1.4%) was supported by a government agency. Interestingly, we failed to identify any research papers that were co-written with authors from water companies.

Figure 9 illustrates the distribution of publications based on venue type. Most publications (54.7%) were published in conferences, 31.2% were published in journals and the remaining 14.1% were published in workshops. Table 3 shows the publication venues for these papers. To answer **RQ4**, the most targeted conference is the World Environmental and Water Resources Congress with 11 papers published; the remaining conference papers were published in a wide range of conferences. The International Workshop on Cyber–Physical Systems for Smart Water Networks, which was established in 2015 and brings together researchers and engineers working on smart water systems, is the most targeted workshop. The most popular journal targeted for publishing security-related papers for water systems is the Journal of Water Resources Planning and Management, published by the American Society of Civil Engineers since the early 1990s. There was not enough data to reliably investigate the role of the conference and journal influencing the publication citations.

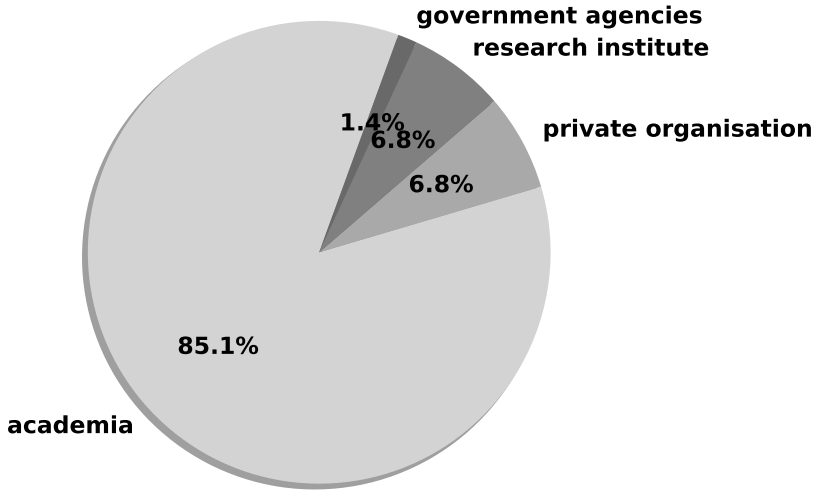

Figure 8. Affiliation of authors.

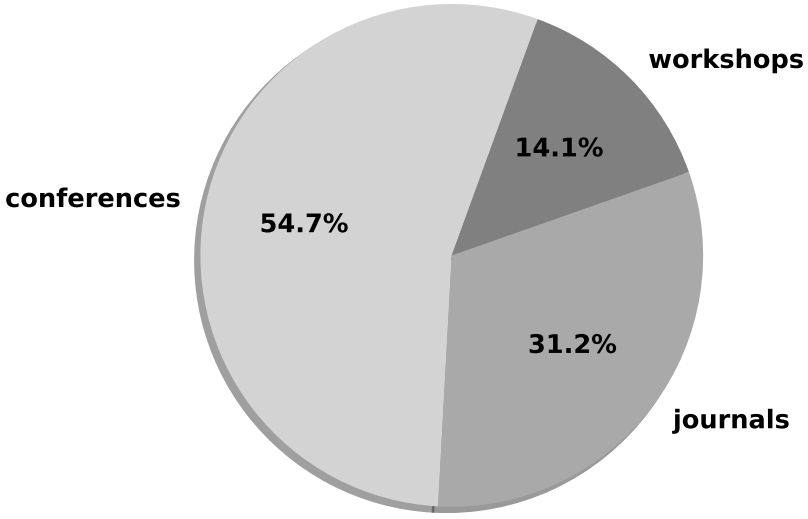

Figure 9. Venues for publication.

Figure 10 shows the results for **RQ5**, the security aspects covered by the publications. Most of the existing work focuses on detection mechanisms. The availability of datasets such as SWaT and WADI [45] has encouraged more research in this area. 31 papers investigated detection models; 10 papers investigating attacks against water systems and

determining their impact; 9 papers on simulation or testbeds; 5 papers used modelling approaches for security analysis; 3 papers developed approaches for risk and resilience management; 2 papers were on datasets; 2 papers covered case studies; 2 papers examined benchmarking; a single paper addressed the development of a security framework; and another paper looked at improving security monitoring capabilities for water systems. In the following sections, we introduce the security aspects covered by the publications and provide a review.

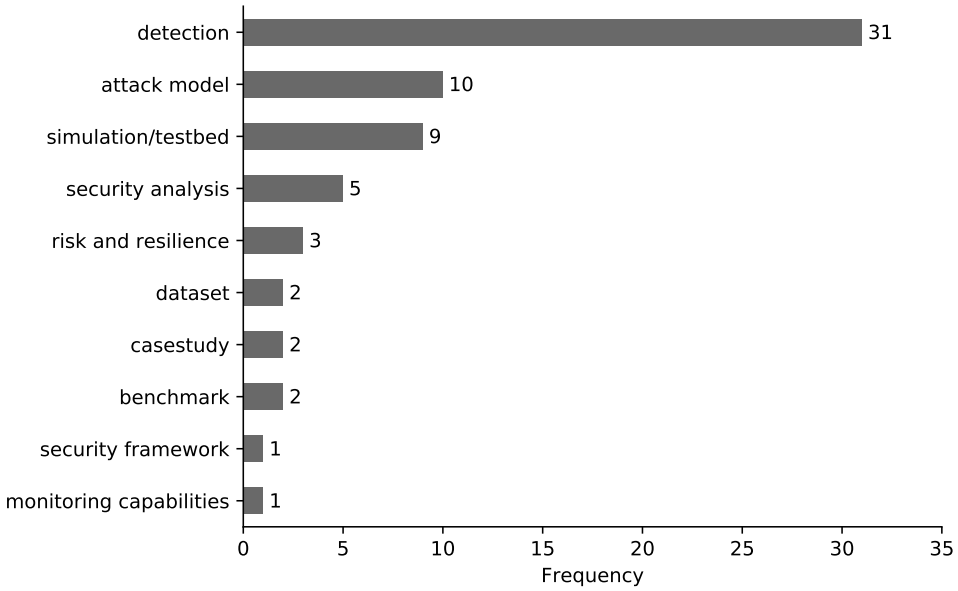

Figure 10. Security aspects covered by publication.

Table 3. Publication venues.

Type	Name	Count
conference	World Environmental and Water Resources Congress	11
workshop	International Workshop on Cyber-Physical Systems for Smart Water Networks	6
journal	Journal of Water Resources Planning and Management	5
journal	Journal of Environmental Engineering	3
conference	IEEE International Conference on Software Quality, Reliability and Security	3
conference	International Conference on Critical Information Infrastructures Security	2
conference	ACM on Asia Conference on Computer and Communications Security	2
journal	IEEE Transactions on Control Systems Technology	2
workshop	International Workshop on the Security of Industrial Control Systems and CPS	1
workshop	International Workshop on Critical Information Infrastructures Security	1
workshop	IEEE/ACM International Workshop on Software Engineering for Smart CPS	1
workshop	ACM Workshop on Cyber–Physical Systems Security and Privacy	1
journal	Water Resources Management	1
journal	Water Research	1
journal	Journal of Systems Science and Systems Engineering	1
journal	International Journal of Critical Infrastructure Protection	1
journal	IEEE Transactions on Dependable and Secure Computing	1
journal	IEEE Signal Processing Magazine	1
journal	IEEE Design and Test	1
journal	Human-centric Computing and Information Services	1
journal	Future Internet	1
journal	Environmental Modelling and Software	1
conference	Pipeline Division Specialty Congress	1
conference	International Symposium on Computer Science and Intelligent Control	1
conference	International Conference on Technology Trends	1
conference	International Conference on Harmony Search Algorithm	1

Table 3. *Cont.*

Type	Name	Count
conference	International Conference on Critical Infrastructure Protection	1
conference	International Conference on Auditory Display	1
conference	IFIP TC 11 International Conference on ICT Systems Security and Privacy Protection	1
conference	IFAC Conference on Cyber–Physical and Human Systems	1
conference	IEEE/ACM Int'l Conference on Cyber, Physical and Social Computing	1
conference	IEEE Pacific Rim International Symposium on Dependable Computing	1
conference	IEEE International Symposium on High Assurance Systems Engineering	1
conference	IEEE International Conference on Machine Learning and Applications	1
conference	IEEE International Conference on Data-Mining Workshops	1
conference	IEEE International Conference on Big Data	1
conference	ACM international conference on Hybrid systems: Computation and Control	1
conference	Annual Computer Security Applications Conference	1

4.2. Classification of Studies

Existing studies were categorised into the following areas: testbeds, simulation and datasets; cyber-attack models; cyber-attacks detection models; model-based security analysis; risk and resilience management; security frameworks; and security benchmarks and case studies. These categories help to answer **RQ6**, showing type of research contributions.

4.2.1. Testbeds, Simulation and Datasets

As it is typically neither possible nor safe to carry out cyber-security research studies that include attacks on operational cyber–physical systems, researchers have been using testbeds and simulation to reproduce the operation and characteristics of real-world systems. A number of testbed and simulation platforms have been proposed for the security of water systems. Table 4 outlines reported tools that have been used to support security research for water systems, including developing datasets for testing intrusion detection and validating mitigation techniques. The most widely known and reputable of these are the Secure Water Treatment (SWaT) testbed [43] and water distribution testbed (WADI) [44], both of which were implemented and deployed at iTrust Centre for research in cyber-security at Singapore University of Technology and Design [46]. SWaT consists of a six-stage water treatment process: raw water processing, chemical dosing, ultrafiltration, water purification (reverse osmosis) and backwashing [46]. The testbed also includes a real layered communication network consisting of layer 0 (sensors, actuators, PLCs) and layer 1 (SCADA, HMI, workstation and historians) of the Purdue model, using both wired and wireless network protocols. The WADI testbed is composed of set of tanks (e.g., reservoir tanks, consumer tanks, raw and returned water tanks), chemical dosing systems, and supporting equipment for water storage and distribution. WADI was designed as an extension to the SWaT [46] testbed and, by combining the capabilities of both testbeds, researchers were able to form a complete and fully functional water treatment, storage and distribution testbed for security research. Both testbeds were designed with international collaborators and engineers from the water sector and the combination has facilitated investigations that include the cascading effects of cyber-attacks between different components of the two testbeds. Researchers have also provided the cyber-security research community with datasets [45] containing normal operation and attack scenarios to allow detection methods to be evaluated. These datasets are multivariate time-series collected from real-time data sources such as sensors and actuators. One of the widely studied datasets in cyber-security research is the SWaT dataset [50] containing normal data streams collected from 51 sensors and actuators collected over 7 days, and attack data consisting of 41 attacks carried out over a period of 4 days. The WADI dataset [45] contains data from 123 sensors and actuators collected over a period of 14 days, and two days with attacks. Given the care in their design and their uniqueness, it is no surprise that a significant amount of research has been carried out using these testbeds and datasets. The iTrust Centre also runs schemes for other local

and international researchers to request access to testbeds, subject to availability and an hourly charge.

Table 4. Testbeds and simulation tools used for cyber-security studies.

Publication	Details	Dataset
WaterBox (2015) [51]	A small-scale cyber–physical testbed designed for an in-lab environment to simulate smart water networks using components designed from acrylic, Arduino boards, small-scale sensors (pressure sensor, flow meter) and a motorised valve (using a small stepper motor).	-
SWaT (2016) [43,46]	An operational small-scale water treatment testbed with real cyber and physical equipment to investigate cyber-security research in 2015 by Singapore University of Technology and Design. It consists of a six-stage water treatment process with the modern-day components.	Available [45,50]
WADI (2016) [44,46]	A testbed launched by Singapore University of Technology and Design funded in 2016 as an extension of SWaT testbed to form a complete water treatment, storage and distribution system.	Available [45]
epanetCPA (2016) [52,53]	EPANET-based toolbox that is designed to assess the impact of cyber–physical attacks.	-
FACIES (2017) [54]	A water distribution system prototype funded by EU project FACIES based on a small fictitious city distributing water to different residential areas with a reservoir represented as tanks of different sizes.	-
RISKNOUGHT (2018) [55–57]	A cyber–physical stress testing platform leveraging EPANET software library to simulate the physical process and a custom network model for SCADA system.	
Water storage control (2018) [58]	A SCADA testbed simulating water storage control consisting of water tank, PLC, historian, HMI, water level sensors and actuators (pumps and valve). The testbed was used to evaluate machine learning detection models against reconnaissance, command injection, and DoS attacks.	-

Other identified testbeds include WaterBox [51], a small-scale cyber-physical testbed designed as an in-lab facility built using Arduino boards, pressure sensors, flow meters, motorised valves, and acrylic structure to simulate smart water networks to carry out experiments related to water systems research including cyber-security and control optimisation. Teixeira et al. [58] developed a SCADA testbed system designed for controlling a water storage tank, simulating the process of water treatment and distribution, to test developed solutions such as machine learning based cyber attack detection models. This testbed includes a PLC (Schneider model M241CE40), HMI, water tanks, water pumps, valves, and sensors for water levels, and uses Modbus communication protocol. Miciolino et al. [54] reports FACIES testbed, emulating a water supply and distribution system for a fictional city to study security of water systems as part of EU project FACIES. The testbed consists of acrylic water tanks, sensors and actuators that are connected to PLCs (Modicon M340, Schneider), a SCADA system and a HMI. The communication protocol used by SCADA and PLC is Modbus over TCP protocol.

Simulation tools developed to study security of water systems include EPANET [59] based tools: epanetCPA [52,53], a simulation toolbox designed for simulating water distribution networks; and RISKNOUGHT (2018) [55–57] developed by STOP-IT project as a cyber-physical stress testing platform for water distribution networks including functionalities to simulate the flow of information between physical (hydraulic model) and cyber layers (SCADA networks).

4.2.2. Cyber-Attack Models

The modelling of attacks is an important part of cyber-security research, because it helps in understanding: the vulnerabilities of cyber–physical systems; the resources required to carry out successful attacks; the impact of attacks; and the resilience of counter-

measures. Over the past decade, attacks against cyber–physical systems have attracted increased interest from the security research community to understand the resources required for attackers to carry out effective attacks.

We identified several papers that developed attack models to examine the behaviour of water systems and the impact of attacks. In [60], researchers investigated stealthy attacks that could cause damage while evading detection. They assumed an attacker with advanced skills and developed resources such as system dynamics, system diagnostic schemes, and the ability to manipulate PV (sensor) data. Attacks were carried out on the Gignac (Southern France) canal network's SCADA system. Researchers were able to design attacks that evaded the diagnostic scheme, which was based on unknown input observers for fault detection and isolation.

Adepu and Mathur [61] investigated single-point cyber-attacks against SWaT testbed and proposed attack detection based on system response to the attacks. Adepu et al. [62] and Tomic et al. [63] investigated jamming attacks against wireless communications in water systems. In [62], researchers carried out attacks against different parts of the SWaT testbed and, in [63], researchers used the Waterbox testbed [51] to investigate the robustness of process control schemes against jamming attacks using different attack strategies. Such attacks have the potential to halt or slow down a process and cause components to fail [62].

Robles-Durazno et al. [64] investigated memory corruption attacks against a PLC used in a water supply process, demonstrating their research using a Festo MPA rig. Researchers investigated memory corruption attacks in three location: attacking PLC inputs by overwriting memory allocated to connected sensors; attacking PLC outputs by overwriting memory for actuators; and attacking PLC working memory, targeting runtime code that contained setpoint variables. Researchers proposed a detection model based on monitoring energy consumption and voltage signals of sensors and actuators. Amin et al. [65] demonstrate stealthy deception attacks against SCADA systems used within water infrastructures.

RISKNOUGHT [55–57] simulation platform developed interaction between physical processes, and the computational and networking layers to simulate a range of cyber–physical threats including cyber-attacks targeting sensors, actuators, PLCs, SCADA and historians, causing physical damage to hydraulic components such as pumps, valves and pipes. Similarly, Taormina et al. [66] included a range of attack scenarios with the epanetCPA [53] toolkit to simulate cyber and physical attacks that target sensors, actuators, PLCs and SCADA, and communication between these components.

Erba et al. [67] investigated adversarial machine learning against ICS used in water distribution systems using WADI and BATADAL datasets. They present two models for concealment attacks to evade detectors that were trained using deep neural networks: (i) a white box attacker that has knowledge of the system and detection model and uses optimisation to generate adversarial samples that are close to the normal operating values of sensors; and (ii) a black box attacker, where the attacker has no knowledge of the detection and uses deep neural networks to learn the behaviour of expected ICS behaviour and produce adversarial sensor readings that resemble real data.

4.2.3. Cyber-Attack Detection Models

Designing effective detection techniques for cyber–physical systems is an important and dynamic area of research. A general list of cyber–physical systems detection models is reported in [68]. In this section, we review models proposed for detecting cyber-attacks in water systems.

A wide variety of approaches have been used to detect abnormal behaviour in water systems. These approaches are illustrated in Table 5. These can be divided into: model-based detection, which tries to model the physical evolution of systems; machine learning models, which learn representative characteristics of a system using data; and statistical models, which use statistical analysis to detect attacks.

Table 5. Papers related to the cyber-attackattack detection.

Publication	Attacks	Application Environment	Dataset	Detection Model
Amin et al. [69]	deception attacks against PVs	a simplified canal hydrodynamic model	-	model-based
Adepu and Mathur [70–73]	bias attacks [74]	SWaT testbed	-	model-based: invariants
Yoong and Heng [75]	-	SWaT testbed	-	machine learning invariants
Miciolino et al. [54]	DoS, replay	FACIES	-	standard deviation
Zohrevand et al. [76]	attacking water flow	water supply system	operational water supply system in Canada	hidden Markov model
Ahmed et al. [77]	false data injection and zero-alarm attacks against PVs and MVs	simulation: EPANET	-	model-based
Moazeni and Khazaei [78]	-	simulation: MATLAB OPTi toolbox	-	model-based: MINLP
Inoue et al. [79]	deception attacks against PVs and MVs	-	SWaT	LSTM and one-class SVM
Hindy et al. [80]	DoS, spoofing	physical testbed	-	classic machine learning methods
Studies using BATADAL dataset [47]	deception attacks, replay against PVs and MVs	-	BATADAL	autoencoders [81,82], MLP and PCA [83,84], data-mining [85,86], NARX [87], rule-based and deep learning [88], model-based (MILP) [89,90], model-based(feature extraction and random forest) [91], PCA, EWMA and RBC [92], ensemble (SOD, LOF and QDA) [93],
Kadosh et al. [94]	deception attacks, replay	C-Town, E-Town WDSs	BATADAL and generated dataset	SVDD
Bakalos et al. [95]	deception attacks against PVs, physical intrusions	water infrastructure SCADA systems	STOP-IT	TDL-CNN
Min et al. [96]	deception attacks against PVs and MVs	simulation: EPANET	-	ANN
Macas et al. [97]	deception attacks against PVs and MVs	-	SWaT	deep autoencoders
Zou et al. [98]	-	WDS in US	-	data-driven estimation (ANNs) and one-class SVM
Ghaeini and Tippenhauer [99]	network attacks	SWaT testbed	-	deep packet inspection

Amin et al. [69] propose a theoretical model-based detection scheme based on hydrodynamic models to detect cyber-attacks against sensor measurements and other anomalous behaviour in canal systems. Adepu and Mathur [70] used the SWaT testbed to detect

cyber-attacks using invariants, the physical conditions that must be true for a process at a given state. Researchers test their approaches using a selection of bias attacks, in which attackers modified sensor outputs and actuator commands by adding a small constant each time [74]. Researchers extended their work in [71,72] to detect bias attacks [74] against sensors and actuators using physics-based invariants for each state of the process, derived from process design for both single-point attacks happening at a single stage, and multiple point attacks that affect multiple sensors and actuators at a single stage [72], and proposed a distributed attack detection method in [73] to detect coordinated cyber-attacks. Yoong and Heng [75] proposed a security framework to develop and evaluate machine learning invariants to detect anomalies, and tested their framework using the SWaT testbed. They used an autoregressive model with exogenous inputs (ARX) combined with group searching to construct machine learning invariants to detect anomalies. The proposed framework is capable of being tested in real-life water treatment plants without causing any disturbances.

Miciolino et al. (2017) [54] proposed a fault detection and network anomaly-based detection models for FACIES testbed by monitoring data generated by sensors and network traffic between PLCs and SCADA which uses Modbus over TCP protocol. Detection uses standard deviation between the normal behaviour and actual observations. Normal behaviour of sensors and network traffic is determined by using statistical averages calculated using data from normal runs.

Zohrevand et al. (2016) [76] used a hidden Markov model (HMM) to design an anomaly-based detection model for a water supply system. Training data was collected from a SCADA-based water supply system in the City of Surrey in British Columbia (Canada) between 2011 and 2014. Working with domain experts, researchers generated anomalous cases and inserted these into the normal data as potential attack data. Four anomalies were constructed by targeting the flow capacity of water: maximum flow, minimum flow, continuous overflow and frequent overflow. Ahmed et al. (2017) [77] used EPANET to simulate a water distribution network to demonstrate a model-based attack detection technique. Detection involves determining the input-output dynamical model of the water distribution network as a set of Linear Time Invariant (LTI) equations. A Kalman Filter is then used to estimate the state of the physical process. The difference between actual measurements and estimations are used to obtain residuals which are then fed into a change detection procedure, CUSUM (cumulative sum control chart) to identify abnormal behaviour. Generated attacks include false data injection (sending modified PVs to controller; and sending false signals to actuators); and controller zero-alarm attack where the attacker changes sensor measurements in such a way that residuals do not cause any alarms. Moazeni and Khazaei [78] proposed a mixed integer nonlinear programming (MINLP) approach to estimate state variables, and tested this on a simulated 6-node water distribution system modelled using the MATLAB OPTi toolbox.

Many machine learning techniques, both supervised and unsupervised, have been used to detect anomalous behaviour. Inoue et al. [79] used a SWaT dataset [50], which consists of 41 cyber and physical attacks [45] against sensors, actuators and controllers including modifying PVs and MVs. Researchers used unsupervised learning approaches from deep learning (long short-term memory neural networks) and one-class support vector machines to detect anomalies.

Hindy et al. [80] built a water system testbed composed of two water tanks, a PLC, a Modicon M238 logic controller, pumps and five sensors that measures various water levels and the presence of water in the tanks. The testbed has two mode of operation, simulating water distribution, and storage. Sensor measurements are sent to the control and monitoring units using the Modbus protocol. Anomalous behaviour is generated as a result of cyber-attacks (DoS, spoofing), system faults and physical attacks (e.g., humans hitting tanks). Classic machine learning algorithms are used to classify anomalous behaviour and affected components using the data gathered and reported by the PLCs. These algorithms are logistic regression, Gaussian naive Bayes, k-nearest neighbors (K-NN), support vector

machine (SVM), decision trees and random forests [80]. They report that the K-NN model achieved the highest accuracy.

Several teams participated in the BATADAL challenge competition [47], developing attack detection for the fictitious C-Town water distribution network (WDN) benchmark [100]. This was built using the epanetCPA water distribution modelling toolkit, and presented at the 2017 World Environmental and Water Resources Congress organized by the Environmental and Water Resources Institute of the American Society of Civil Engineers (EWRI/ASCE). Three datasets [45], one with normal operational data, and two datasets (one for training, one for testing) containing cyber-attacks, were given to each competing team. Generated cyber-attacks were deception attacks (against PVs and MVs and SCADA data) and replay attacks. Taormina and Galelli [81,82] used autoencoders (deep neural networks) in detecting attacks. Abokifa et al. [83,84] proposed a detection approach composed of three layers to detect anomalies in the BATADAL datasets; first removing outliers using statistical analysis then, using a feed forward artificial neural network (ANN), a multilayer perceptron (MLP) to identify anomalies and, finally, principal component analysis (PCA) to identify multiple affected sensors. Giacomoni et al. [85] developed two detection approaches based on data-mining. The first of these is a method using actuator rules to ensure readings from the SCADA are within defined normal ranges. The second method uses an optimization routine that extracts low-dimensionality components of the data, and thereby separates normal operation data from attack data. Pasha et al. [86,101] also used a data-mining approach on BATADAL datasets based on extracting control rules, pattern recognition, PCA, and relationship between hydraulic and system parameters. Brentan et al. [87] applied autoregressive networks with exogenous inputs (NARX), a recurrent neural network. Housh and Ohar [89,90] used physical simulation to model the system to detect cyber-attacks. Their model-based approach uses mixed integer linear programming (MILP) to estimate the hydraulic processes of the water distribution systems under normal operating conditions to produce expected errors between the actual measurements and estimated model. The difference between the expected and actual value is used to detect attacks. Chandy et al. [88] developed an ensemble model comprising two models to detect attacks for the BATADAL detection challenge competition. The first uses physical and operational rules and violations to generate events. The second uses these events along with raw data to train a deep learning model, a convolutional variational autoencoder, to detect attacks. Aghashahi et al. [91] first extracted features related to the characteristics of the attack and no-attack data by using a covariance matrix and distance measure of every data point. Then, a random forest classifier was used to classify these characteristics as attack and normal operation. A detailed description of the competition and a discussion of results can be found in [47]. MarcosQuiñones-Grueiro [92] combined widely used signal processing techniques, PCA, the adaptive exponential weighted moving average chart (EWMA) and the reconstruction-based contribution (RBC) method to detect attacks and to diagnose the area of the network that was under attack using the BATADAL dataset. Ramotsoela et al. [93] used the BATADAL dataset to evaluate some of the traditional anomaly detection approaches to detect attacks in WDS, and proposed an ensemble technique. The proposed ensemble technique combines the subspace outlier degree (SOD) algorithm, a distance-based shared nearest neighbors approach designed to detect outliers in high-dimensional data [102] and a local outlier factor (LOF) algorithm [103] to detect outliers in low-dimensional data. Both algorithms are run in parallel for each predicted datapoint and feed their outputs to a quadratic discriminant analysis (QDA) process to classify datapoints into anomalous or normal. Kadosh et al. [94] used a support vector data description (SVDD) classifier to propose a one-class cyber-attack detection model to detect attacks in WSD using both the BATADAL dataset and epanetCPA.

Bakalos et al. [95] developed a cyber-attack detection approach for water systems using multimodal data fusion and adaptive deep learning. Multimodal data fusion involves combining different channels of information, including visual data from thermal camera streams, Wi-Fi reflection, and ICS data. The weight attached to each of these streams of

data is determined through a deep learning model process. The proposed adaptive deep learning approach uses a tapped delay line (TDL) convolutional neural network (CNN) with autoregressive moving average [95]. The data used to evaluate the approach is from STOP-IT project.

Min et al. [96] used an artificial neural network to detect attacks against a water distribution network using the EPANET simulator [84]. Macas et al. [97] used an "unsupervised attention-based spatio-temporal autoencoder for anomaly detection (STAE-AD)" model to detect attacks against water infrastructures using the SWaT dataset. Zou et al. [98] proposed a hybrid model making use of an MLP and a one-class SVM. MLP was used to forecast measurement parameters, and prediction errors were used to train a one-class SVM to classify outliers; finally, Bayesian sequence analysis was used to detect contamination attacks against water distribution systems.

Majority of cyber-attack detection models reviewed focus on detecting anomalous behaviour by monitoring and analyzing physical process variables, and failed to monitor industrial control network traffic and use this knowledge to detect cyber-attacks. Ghaeini and Tippenhauer [99] proposed a hierarchical monitoring intrusion detection system (HAMIDS) for ICS to collect network events in different layers of industrial networks. HAMIDS extends the Bro, an open-source tool for monitoring and analyzing network traffic. IDS sensors are installed on different layers of industrial networks to monitor network events. These events are then aggregated and processed in a central cluster to detect malicious behaviour. HAMIDS was validated using a range of network attacks (e.g., ARP poising, network flooding and man in the middle attacks) against SWaT testbed.

Proposed detection approaches are evaluated for effectiveness using (i) operational data from real-world systems; (ii) testbeds; and (iii) simulation. Existing studies show a wide variety of techniques that were applied to detect cyber-attacks against water systems; however, making a reliable comparison among detection approaches is not feasible due to a lack of common performance metrics and/or missing reported performance data, different datasets and sizes.

4.3. Model-Based Security Analysis

Several research studies focused on using modelling approaches to analyse the security of water systems and to identify vulnerabilities.

Kang et al. [104] proposed a model-based security analysis for a water treatment system. Testing their approach on SWaT, they modelled the interaction between the physical plant and controller using approximate, discrete models to discover and explore potential attacks. The model is constructed using a first-order modelling language Alloy to capture, as state transition rules, connections among various components and the behaviour of the plant.

Motivated by malware techniques that hide critical information from operators while executing an attack (e.g., Stuxnet), Patloll et al. [105] proposed a multiple security domain non-deducibility (MSDND) model [106] using belief, information transfer and trust (BIT) logic [107] to identify critical information that attackers may hide. BIT logic is used to reason about the reliability of data moving between entities, defined as the belief and trust one entity has in information received from another entity. A system is decomposed into components, and each component that could change the state of the state is treated as a separate domain. Requiring development of invariants, an information execution flow across these domains starting from source to destination is monitored to identify when vulnerabilities that have been exploited have resulted in invariant violation. Mishra et al. [108] proposed an agent-based modelling framework to model critical CPS and their interdependencies, to understand the impact of attacks on interconnected critical infrastructures; they evaluated the application of the model to a water distribution system and used invariant-based method [70] to generate rules to detect attacks.

Taormina et al. [66] and Hunter et al. [109] proposed a modelling approach to quantify the hydraulic behaviour of the system (such as tank overflow, variation in pumps) under

cyber–physical attacks by defining components of a system, and specifying attack variables (starting time, duration). They give simulation results using the epanetCPA toolbox and the C-Town network [100].

4.4. Risk and Resilience Management

A small number of studies worked on methods to support risk and resilience management.

Moraitis et al. [110] describes a methodology to quantify the impact of cyber–physical attacks on water distribution networks. The methodology is based on quantifying failures described under categories (magnitude, propagation, severity, crest factor, rapidity) against user-defined service levels. A proposed model is demonstrated using the C-Town WDN.

Jeong [111] discusses the development of a risk management framework for water infrastructure against intentional attacks, including cyber-attacks based on vulnerability assessment and consequence assessment of attacks. The proposed vulnerability assessment involves the development of a hierarchical structure of the system to identify all water infrastructure components, using expert knowledge and fuzzy hierarchical analysis. The recommended consequence assessment is based on the time to restore the system to its normal operation, and the areas affected by the attack, and the expected damage is based on attacker's and defender's capabilities.

Shin et al. [112] investigated resilience strategies against water CPS. Resilience is characterized in terms of four capabilities [112]: (i) ability to withstand disruption; (ii) absorptive capability (if disruption is unavoidable then minimize undesirable consequences; (iii) adaptive capability (adjusting to disrupted and undesirable conditions); (iv) restorative capability (recover quickly to completely normal operation). A resilience metric is proposed to measure the resilience of water systems against cyber-attack, and the C-town benchmark water distribution system is used as a case study to demonstrate the proposed metric.

4.5. Security Frameworks

Modern water treatment infrastructures consist of interconnected systems layered in a hierarchy, such as a supervisory layer consisting of SCADA systems, and a control layer composed of PLCs, sensors and actuators. Data flows occur between these layers via multiple communication networks. Mathur [113] proposes a multilayer security framework composed of seven layers of countermeasures applied to different network layers to secure water treatment systems. Proposed countermeasures include attack prevention mechanisms (firewalls), attack detection mechanisms (intrusion detection systems, process anomaly detection), and post-attack mechanisms that could bring the process back to a normal or manageable state. A partial implementation of the proposed framework was tested on the SWaT testbed.

4.6. Security Benchmarks and Case Studies

TNO (Netherlands Organisation for Applied Scientific Research—an independent research organisation) and the NICC (the Netherlands Infrastructure Cybercrime unit), carried out a study [114] to understand the current state of cyber-security of process control in the drinking water sector in the Netherlands. Researchers report that a large variance of security posture was found among organisations; the data collected exposed serious weaknesses in each company. As the study contained sensitive national data, confidentiality of the organisations was maintained and the reported analyses were based on artificially aggregated data. The study was effective and resulted in the development of good practices for SCADA security for drinking water organisations, which are available both in Dutch and English [115]. Building on this work, Burghouwt et al. [116] measured the cyber-security state of the 19 water management organisations in the Netherlands through an improved questionnaire. Researchers identified a lack of uniformity on security postures between organisations, partly due to ineffective management of security responsibilities. They designed and built DESI [116], a simulator to demonstrate cyber–physical attack scenarios and improve cyber-attack knowledge.

A case study paper was presented in [117] investigating access control mechanisms in industrial control systems conducted on the WADI testbed, to show how the lack of effective access control could lead to malicious behaviour. Researchers revealed that a lack of access control in network protocols, systems and field devices used in ICS is making these systems vulnerable to attacks.

A critical case study for security of water systems is the Marooch water breach incident. Slay and Miller [29] discusses this incident and reports the lessons learned from the incident emphasising the need for effective, reliable and economically viable security countermeasures including intrusion detection systems for SCADA networks, better management of security policies and procedures, investment in security training for staff, and a wider and sustainable collaboration between academia, industry, vendors and government agencies to tackle existing and future security threats.

4.7. Security Monitoring Capabilities

One of the papers identified dealt with improving security monitoring capabilities for water distribution systems. In [118], researchers propose sonification, data in audio, to help system operators avoid cognitive overloading with visual information to raise alarms for cyber-attacks on water distribution systems. Motivated by prior work on sonification, designed to improve monitoring capabilities, researchers designed a sonification system to reduce the overload of human operators faced with visual channels, to support better decision-making for a water facility by sonifying the outputs of an anomaly detection model. Current anomaly detection models are represented as visual diagrams showing anomalous data points at a given time and often an alarm is raised when a threshold is reached.

5. Open Issues and Future Research Areas

Results obtained from the systematic review show that research has made a significant contribution to the security of water systems. In the following sections we discuss limitations of existing studies and highlight some areas for future research.

5.1. Building Testbeds for Water Systems

Much of the existing research in this area involves a pool of resources (SWaT and WADI testbeds, epanetCPA toolbox, and datasets) provided by the iTrust Centre for research in cyber-security. Researchers associated with the iTrust Centre demonstrate the importance of developing and providing access to a real physical testbed for carrying out security research. Most of the existing studies have focused on drinking water systems, primarily those responsible for water distribution. Given the diversity of water and wastewater systems, more work in this area would provide obvious benefits, especially through testbeds involving water systems such as sewer and wastewater systems, and irrigation systems; these could be used to further validate the applicability of existing research. Although of immense value, building and maintaining realistic operational testbeds is not an easy task and requires significant and ongoing access to resources, skills and people.

5.2. Threat and Attack Models

Existing attack models primarily make use of manual and single-point attacks targeting single measurement variables (sensor readings) or control commands. However, stealthy attacks, those trying to cause damage and at the same time remain undetected, may necessitate multi-point attacks if they are to evade detection mechanisms and operators. This area is starting to receive increased attention from researchers investigating the security of CPS [67]; however, more effort is required to understand how these attacks can be performed and what the limits on their effectiveness might be. Consequently, few studies have verified the effectiveness of existing detection models against these attacks.

5.3. Attack Detection Models

Many studies designed to detect attacks against water CPS use machine learning-based anomaly detection models, in which normal operational data is the primary (or sole) resource as there is often insufficient anomalous data to create models using supervised approaches. It is not readily possible to compare the performance of existing detection models, or to determine their generality or the reproducibility of their results. This is due both to a lack of datasets, leading to poor diversity in assumptions and plant models, and to a lack of common performance metrics. Where common datasets and performance metrics have been used, as in the case of, say, the SWaT and WADI datasets, reported results suggest that deep learning-based anomaly detection models perform better than conventional anomaly detection models. However, further studies are required to build confidence that such performance improvement is real.

As is usually the case with intrusion detection studies for CPS, the effectiveness of the proposed solutions were measured using conventional performance metrics, including accuracy, precision, recall, F-score, false positives and false negatives. These performance metrics were not designed for multivariate time-series datasets of CPS, in which anomalies usually occur in bursts [119]. Even when using these conventional performance metrics, some fail to report false positives and none of the studies reported detection latency, which is an important metric for critical systems [68] as early detection is critical for CPS.

Over the last decade, there has been an increase in number of CPS applying deep learning models to detect anomalous behaviour and datasets such as BATADAL, SWaT and WADi have contributed to some of these studies. However, studies from other fields have shown that machine learning-based approaches are rather vulnerable to accidental or intentional corruption of training data sets; thus, say, adversarial attacks can influence detection outcomes [120]. At the same time, there is a significant number of research studies that focus on improving the robustness of such models [121]. At present, however, such work is invariably targeted at other fields of study, most notably computer vision, and we are yet to understand the possible risks in the application of learning models to CPS.

The generation of attack or anomalous behaviour for testing detection models is often done manually. Typically, measurement values or control signals are modified, and performance data is collected both with and without these variations. However, such an approach assumes that the modifications are representative of those that will be experienced in reality, and this assumption is tenuous at best. Furthermore, over time, CPS actuators and sensors degrade as a result of ageing and become more prone to noise. As a result, normal behaviour is itself non-stationary and it will be necessary either to use richer training sets and models that capture temporal change, or to use online learning. The latter is again vulnerable to changes induced by an adversary that are intended to pervert the detection mechanism. There is therefore a pressing need to increase the attention paid to the practicalities associated with actual deployment, including the usability and maintainability of proposed detection models.

Identifying anomalous behaviour should ideally be followed by the raising of an alert that identifies the potential cause and so determines a strategy to be followed for mitigation. However, existing studies often stop at detection. Future work is therefore required to investigate approaches that identify the root cause of anomalous behaviour, locate compromised devices and respond and mitigate further damage in a timely manner.

5.4. Collaboration with Industry

Although several studies have consulted with engineers who have experience in dealing with water systems, we failed to identify any publications that were written by industry. There is currently a lack of collaborative work between industry and academia in this area. Securing water systems requires a multidisciplinary effort that involves both the designers and operators of these systems and academics working at the leading edge of technology to ensure that security research pushes the boundaries of the possible while remaining applicable and usable.

6. Conclusions

In this paper, we have systematically reviewed the existing peer-reviewed research efforts to secure water systems, and have identified limitations in those research efforts and possible future directions for securing next generation of smart water CPS. This study provides guidance for understanding the existing security research for developing secure smart water systems.

In comparison to other utilities such as electricity, the security of water systems has not received much research attention in the past, but this is changing, and there has been an increase in the number of studies since 2016 supported by EC research and innovation funding programs and international funding opportunities. The studies reviewed in this paper are encouraging, but they require further work for validation and deployment on real water systems. Most of the existing studies, including testbeds, simulation tools and datasets, have focused on drinking water treatment, supply and distribution. Further studies are required to build testbeds, simulation and datasets that investigate security of non-drinking water sectors such as wastewater treatment systems, stormwater management and systems for agriculture and irrigation.

Finally, development of a comprehensive usable security framework that covers different aspects of security, from prevention to detection, response and mitigation requires a multidisciplinary approach involving academia-industry-government cooperation.

Author Contributions: Conceptualization, N.T., S.H. and J.W.; methodology, N.T. and S.H.; data curation, N.T.; writing—original draft preparation, N.T.; writing—review and editing, N.T., P.H., S.H. and J.W.; visualization, N.T. and P.H.; supervision, S.H. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the UK Engineering and Physical Sciences Research Council (EPSRC), under The PETRAS National Centre of Excellence for IoT Systems Cybersecurity, grant number EP/S035362/1.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. WWAP (United Nations World Water Assessment Programme)/UN-Water The United Nations World Water Development Report 2018: Nature-based Solution for Water. Paris, UNESCO. 2018. Available online: www.unwater.org/publications/world-water-development-report-2018/ (accessed on 6 November 2020).
2. Bank, W. The World Bank and the International Water Association to Establish a Partnership to Reduce Water Losses. 2016. Available online: https://www.worldbank.org/en/news/press-release/2016/09/01/the-world-bank-and-the-international-water-association-to-establish-a-partnership-to-reduce-water-losses (accessed on 6 November 2020).
3. Discoverwater. Leaking Pipes. 2020. Available online: https://discoverwater.co.uk/leaking-pipes (accessed on 14 November 2020).
4. Li, J.; Yang, X.; Sitzenfrei, R. Rethinking the Framework of Smart Water System: A Review. *Water* **2020**, *12*, 412. [CrossRef]
5. Giudicianni, C.; Herrera, M.; Nardo, A.D.; Adeyeye, K.; Ramos, H.M. Overview of Energy Management and Leakage Control Systems for Smart Water Grids and Digital Water. *Modelling* **2020**, *1*, 134–155. [CrossRef]
6. Adedeji, K.B.; Hamam, Y. Cyber-Physical Systems for Water Supply Network Management: Basics, Challenges, and Roadmap. *Sustainability* **2020**, *12*, 9555. [CrossRef]
7. Ofwat. *PR19 Draft Determinations: UK Government Priorities 2019 Price Review Draft Determinations*; Technical Report; Ofwat: Birmingham, UK, 2019.
8. Ofwat. *Time to Act, Together: Ofwat's Strategy*; Technical Report; Ofwat: Birmingham, UK, 2019.
9. Schickhuber, G.; McCarthy, O. Distributed fieldbus and control network systems. *Comput. Control Eng. J.* **1997**, *8*, 21–32. [CrossRef]
10. SWAN Forum. A Layered View of Smart Water Networks. Available online: https://www.swan-forum.com/swan-tools/a-layered-view (accessed on 1 November 2020).

11. Falliere, N.; Murchu, L.O.; Chien, E. *W32.Stuxnet Dossier (Version 1.4)*; White Paper, Symantec Security Response; Symantec: Mountain View, CA, USA, 2011.
12. Symantec. *W32.Duqu: The Precursor to the Next Stuxnet (Version 1.4)*; White Paper, Symantec Security Response; Symantec: Mountain View, CA, USA, 2011.
13. Kaspersky. BlackEnergy APT Attacks in Ukraine. Available online: https://www.kaspersky.co.uk/resource-center/threats/blackenergy (accessed on 30 November 2020).
14. Havex Hunts For ICS/SCADA Systems. 2014. Available online: https://www.f-secure.com/weblog/archives/00002718.html (accessed on 30 October 2020).
15. Gleick, P.H. Water and terrorism. *Water Policy* **2006**, *8*, 481–503. [CrossRef]
16. Interpol. The Protection of Critical Infrastructure against Terrorist Attacks: Compendium of Good Practices. Compiled by CTED and UNOCT in 2018. 2018. Available online: https://www.un.org/sc/ctc/wp-content/uploads/2019/01/Compendium_of_Good_Practices_Compressed.pdf (accessed on 1 August 2020).
17. Hassanzadeh, A.; Rasekh, A.; Galelli, S.; Aghashahi, M.; Taormina, R.; Ostfeld, A.; Banks, M.K. A Review of Cybersecurity Incidents in the Water Sector. *J. Environ. Eng.* **2020**, *146*, 03120003. [CrossRef]
18. Clark, R.M.; Panguluri, S.; Nelson, T.D.; Wyman, R.P. Protecting Drinking Water Utilities from Cyberthreats. *J. AWWA* **2017**, *109*, 50–58. [CrossRef]
19. ZDNet. Israel Government Tells Water Treatment Companies to Change Passwords. 2020. Available online: https://www.zdnet.com/article/israel-says-hackers-are-targeting-its-water-supply-and-treatment-utilities/ (accessed on 6 November 2020).
20. The Coloradoan. Cyberattacker Demands Ransom from Northern Colorado Utility. 2019. Available online: https://eu.coloradoan.com/story/money/2019/03/14/cyberattacker-demands-ransom-colorado-utility/3148951002/ (accessed on 11 September 2020).
21. Eweek. Water Utility in Europe Hit by Cryptocurrency Malware Mining Attack. 2018. Available online: https://www.eweek.com/security/water-utility-in-europe-hit-by-cryptocurrency-malware-mining-attack (accessed on 11 September 2020).
22. The Registry. Water Treatment Plant Hacked, Chemical Mix Changed for Tap Supplies. 2016. Available online: https://www.theregister.com/2016/03/24/water_utility_hacked (accessed on 14 November 2020).
23. The New York Times. A Dam, Small and Unsung, Is Caught Up in an Iranian Hacking Case. 2016. Available online: https://www.nytimes.com/2016/03/26/nyregion/rye-brook-dam-caught-in-computer-hacking-case.html (accessed on 11 September 2020).
24. The United States Department of Justice. United States District Court Southern District of New York: Sealed Indictment. 2016. Available online: https://www.justice.gov/opa/file/834996/download (accessed on 31 December 2020).
25. Govtech. Report: Hacking Lands Florida Wastewater Official in Hot Water. 2012. Available online: https://www.govtech.com/public-safety/Report-Hacking-Lands-Florida-Wastewater-Official-in-Hot-Water.html (accessed on 31 December 2020).
26. Computer World. Insider charged with hacking California canal system. 2007. Available online: https://www.computerworld.com/article/2540235/insider-charged-with-hacking-california-canal-system.html (accessed on 10 October 2020).
27. TechRepublic. Pennsylvania Water System Hack Demonstrates Lax Security. 2006. Available online: https://www.techrepublic.com/blog/it-security/pennsylvania-water-system-hack-demonstrates-lax-security/ (accessed on 11 September 2020).
28. The MITRE Corporation. Malicious Control System Cyber Security Attack Case Study–Maroochy Water Services, Australia. 2008. Available online: http://www.mitre.org/sites/default/files/pdf/08_1145.pdf (accessed on 11 September 2020).
29. Jill, J.S.; Miller, M. Lessons Learned from the Maroochy Water Breach. In *Critical Infrastructure Protection*; Goetz, E., Shenoi, S., Eds.; Springer: New York, NY, USA, 2008; Volume 253, pp. 73–82.
30. Rid, T.; Buchanan, B. Attributing Cyber Attacks. *J. Strateg. Stud.* **2015**, *38*, 4–37. [CrossRef]
31. Rogers, M.K. A two-dimensional circumplex approach to the development of a hacker taxonomy. *Digit. Investig.* **2006**, *3*, 97–102. [CrossRef]
32. Green, B.; Krotofil, M.; Abbasi, A. On the Significance of Process Comprehension for Conducting Targeted ICS Attacks. In Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and PrivaCy, CPS '17, Dallas, TX, USA, 3 November 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 57–67. [CrossRef]
33. Dragos. Cyber Threat Perspective Manufacturing Sector. 2020. Available online: https://www.dragos.com/resource/manufacturing-threat-perspective/ (accessed on 5 December 2020).
34. Gill, H. From Vision to Reality: Cyber-Physical Systems. In Proceedings of the HCSS National Workshop on New Research Directions for High Confidence Transportation CPS: Automotive, Aviation, and Rail, Washington, DC, USA, 18–20 November 2008.
35. Lee, E.A. The Past, Present and Future of Cyber-Physical Systems: A Focus on Models. *Sensors* **2015**, *15*, 4837–4869. [CrossRef] [PubMed]
36. Stouffer, K.; Zimmerman, S.; Timothy, T.C.; Lubell, J.; Cichonski, J.; McCarthy, J. *NISTIR 8183: Cybersecurity Framework Manufacturing Profile*; Technical Report; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2017.
37. Hu, V.; Ferraiolo, D.; Kuhn, R. *Assessment of Access Control Systems*; Technical Report; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2006.
38. Initiative, J.T.F.T. *Security and Privacy Controls for Federal Information Systems and Organizations, NIST Special Publication 800-53 Revision 4*; Technical Report; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2015.

39. Ross, R.; McEvilley, M.; Oren, C.J. *Systems Security Engineering Considerations for a Multidisciplinary Approach in the Engineering of Trustworthy Secure Systems, NIST Special Publication 800-160*; Technical Report; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2016.
40. Stouffer, K.; Lightman, S.; Pillitteri, V.; Abrams, M.; Hahn, A. *NIST Special Publication 800-82: Guide to Industrial Control Systems (ICS) Security*; Technical Report; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2014.
41. Hahn, A.; Thomas, R.K.; Lozano, I.; Cardenas, A. A multi-layered and kill-chain based security analysis framework for cyber-physical systems. *Int. J. Crit. Infrastruct. Prot.* **2015**, *11*, 39–50. [CrossRef]
42. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.; The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med.* **2009**, 6, 1–6. [CrossRef] [PubMed]
43. Mathur, A.P.; Tippenhauer, N.O. SWaT: A water treatment testbed for research and training on ICS security. In Proceedings of the 2016 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater), Vienna, Austria, 11 April 2016; pp. 31–36.
44. Ahmed, C.M.; Palleti, V.R.; Mathur, A.P. WADI: A Water Distribution Testbed for Research in the Design of Secure Cyber Physical Systems. In Proceedings of the 3rd International Workshop on Cyber-Physical Systems for Smart Water Networks, Pittsburgh, PA, USA, 18–21 April 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 25–28. [CrossRef]
45. ITrust. Dataset Characteristics: SWaT, WADI and BATADAL. Available online: https://itrust.sutd.edu.sg/itrust-labs_datasets/dataset_info/ (accessed on 8 November 2020).
46. iTrust—Singapore University of Technology and Design (SUTD). Testbeds. Available online: https://itrust.sutd.edu.sg/testbeds (accessed on 30 November 2020).
47. Taormina, R.; Galelli, S.; Tippenhauer, N.O.; Salomons, E.; Ostfeld, A.; Eliades, D.G.; Aghashahi, M.; Sundararajan, R.; Pourahmadi, M.; Banks, M.K.; et al. Battle of the Attack Detection Algorithms: Disclosing Cyber Attacks on Water Distribution Networks. *J. Water Resour. Plan. Manag.* **2018**, *144*, 04018048. [CrossRef]
48. Facies Project. Available online: http://facies.dia.uniroma3.it/ (accessed on 30 November 2020).
49. The STOP-IT Project. Available online: https://stop-it-project.eu/ (accessed on 30 November 2020).
50. Goh, J.; Adepu, S.; Junejo, K.N.; Mathur, A. A Dataset to Support Research in the Design of Secure Water Treatment Systems. In Proceedings of the International Conference on Critical Information Infrastructures Security, Paris, France, 10–12 October 2017; Havarneanu, G., Setola, R., Nassopoulos, H., Wolthusen, S., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 88–99.
51. Kartakis, S.; Abraham, E.; McCann, J.A. WaterBox: A Testbed for Monitoring and Controlling Smart Water Networks. In Proceedings of the 1st ACM International Workshop on Cyber-Physical Systems for Smart Water Networks, CySWater'15, Seattle, WA, USA, 14–16 April 2015; Association for Computing Machinery: New York, NY, USA, 2015. [CrossRef]
52. Taormina, R.; Galelli, S.; Tippenhauer, N.; Ostfeld, A.; Salomons, E. Assessing the Effect of Cyber-Physical Attacks on Water Distribution Systems. In Proceedings of the World Environmental and Water Resources Congress 2016, Palm Beach, FL, USA, 22–26 May 2016; pp. 436–442. [CrossRef]
53. Taormina, R.; Galelli, S.; Douglas, H.; Tippenhauer, N.; Salomons, E.; Ostfeld, A. A toolbox for assessing the impacts of cyber-physical attacks on water distribution systems. *Environ. Model. Softw.* **2019**, *112*, 46–51. [CrossRef]
54. Etchevés Miciolino, E.; Setola, R.; Bernieri, G.; Panzieri, S.; Pascucci, F.; Polycarpou, M.M. Fault Diagnosis and Network Anomaly Detection in Water Infrastructures. *IEEE Des. Test* **2017**, *34*, 44–51. [CrossRef]
55. Nikolopoulos, D.; Makropoulos, C.; Kalogeras, D.; Monokrousou, K.; Tsoukalas, I. Developing a Stress-Testing Platform for Cyber-Physical Water Infrastructure. In Proceedings of the 2018 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater), Porto, Portugal, 10–13 April 2018; pp. 9–11. [CrossRef]
56. Nikolopoulos, D.; Moraitis, G.; Bouziotas, D.; Lykou, A.; Karavokiros, G.; Makropoulos, C. RISKNOUGHT: A cyber-physical stress-testing platform for water distribution networks. In Proceedings of the 11th World Congress on Water Resources and Environment (EWRA 2019) Managing Water Resources for a Sustainable Future, Madrid, Spain, 25–29 June 2019.
57. Nikolopoulos, D.; Moraitis, G.; Bouziotas, D.; Lykou, A.; Karavokiros, G.; Makropoulos, C. Cyber-Physical Stress-Testing Platform for Water Distribution Networks. *J. Environ. Eng.* **2020**, *146*, 04020061. [CrossRef]
58. Teixeira, M.; Salman, T.; Zolanvari, M.; Jain, R.; Meskin, N.; Samaka, M. SCADA System Testbed for Cybersecurity Research Using Machine Learning Approach. *Future Internet* **2018**, *10*, 76. [CrossRef]
59. EPANET Application for Modeling Drinking Water Distribution Systems. United States Environmental Protection Agency. Available online: https://www.epa.gov/water-research/epanet (accessed on 31 December 2012).
60. Amin, S.; Litrico, X.; Sastry, S.; Bayen, A.M. Cyber Security of Water SCADA Systems—Part I: Analysis and Experimentation of Stealthy Deception Attacks. *IEEE Trans. Control Syst. Technol.* **2013**, *21*, 1963–1970. [CrossRef]
61. Adepu, S.; Mathur, A. An Investigation into the Response of a Water Treatment System to Cyber Attacks. In Proceedings of the 2016 IEEE 17th International Symposium on High Assurance Systems Engineering (HASE), Orlando, FL, USA, 7–9 January 2016; pp. 141–148.
62. Adepu, S.; Prakash, J.; Mathur, A. WaterJam: An Experimental Case Study of Jamming Attacks on a Water Treatment System. In Proceedings of the 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), Prague, Czech Republic, 25–29 July 2017; pp. 341–347.

63. Tomić, I.; Breza, M.J.; Jackson, G.; Bhatia, L.; McCann, J.A. Design and Evaluation of Jamming Resilient Cyber-Physical Systems. In Proceedings of the 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Halifax, NS, Canada, 30 July–3 August 2018; pp. 687–694.
64. Robles-Durazno, A.; Moradpoor, N.; McWhinnie, J.; Russell, G.; Maneru-Marin, I. *Implementation and Detection of Novel Attacks to the PLC Memory of a Clean Water Supply System*; Botto-Tobar, M., Pizarro, G., Zúñiga-Prieto, M., D'Armas, M., Zúñiga Sánchez, M., Eds.; Technology Trends; Springer International Publishing: Cham, Switzerland, 2019; pp. 91–103.
65. Amin, S.; Litrico, X.; Sastry, S.S.; Bayen, A.M. Stealthy Deception Attacks on Water SCADA Systems. In Proceedings of the 13th ACM International Conference on Hybrid Systems: Computation and Control, HSCC '10, Stockholm, Sweden, 12–15 April 2010; Association for Computing Machinery: New York, NY, USA, 2010; pp. 161–170. [CrossRef]
66. Taormina, R.; Galelli, S.; Tippenhauer, N.O.; Salomons, E.; Ostfeld, A. Characterizing Cyber-Physical Attacks on Water Distribution Systems. *J. Water Resour. Plan. Manag.* **2017**, *143*, 04017009. [CrossRef]
67. Erba, A.; Taormina, R.; Galelli, S.; Pogliani, M.; Carminati, M.; Zanero, S.; Tippenhauer, N.O. Constrained Concealment Attacks against Reconstruction-Based Anomaly Detectors in Industrial Control Systems. In Proceedings of the Annual Computer Security Applications Conference, ACSAC '20, Austin, TX, USA, 7–10 December 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 480–495. [CrossRef]
68. Mitchell, R.; Chen, I.R. A Survey of Intrusion Detection Techniques for Cyber-Physical Systems. *ACM Comput. Surv.* **2014**, *46*. [CrossRef]
69. Amin, S.; Litrico, X.; Sastry, S.S.; Bayen, A.M. Cyber Security of Water SCADA Systems—Part II: Attack Detection Using Enhanced Hydrodynamic Models. *IEEE Trans. Control Syst. Technol.* **2013**, *21*, 1679–1693. [CrossRef]
70. Adepu, S.; Mathur, A. Using Process Invariants to Detect Cyber Attacks on a Water Treatment System. In Proceedings of the ICT Systems Security and Privacy Protection, Ghent, Belgium, 30 May–1 June 2016; Hoepman, J.H., Katzenbeisser, S., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 91–104.
71. Adepu, S.; Mathur, A. Distributed Detection of Single-Stage Multipoint Cyber Attacks in a Water Treatment Plant. In Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, ASIA CCS '16, Xi'an, China, 30 May 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 449–460. [CrossRef]
72. Adepu, S.; Mathur, A. From Design to Invariants: Detecting Attacks on Cyber Physical Systems. In Proceedings of the 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), Prague, Czech Republic, 17 July 2017; pp. 533–540.
73. Adepu, S.; Mathur, A. Distributed Attack Detection in a Water Treatment Plant: Method and Case Study. *IEEE Trans. Dependable Secur. Comput.* **2018**. [CrossRef]
74. Cárdenas, A.A.; Amin, S.; Lin, Z.S.; Huang, Y.L.; Huang, C.Y.; Sastry, S. Attacks against Process Control Systems: Risk Assessment, Detection, and Response. In Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security, ASIACCS '11, Hong Kong, China, March, 20–21 March 2011; Association for Computing Machinery: New York, NY, USA, 2011; pp. 355–366. [CrossRef]
75. Yoong, C.H.; Heng, J. Framework for Continuous System Security Protection in SWaT. In Proceedings of the 2019 3rd International Symposium on Computer Science and Intelligent Control, ISCSIC 2019, Amsterdam, The Netherlands, 25–27 September 2019; Association for Computing Machinery: New York, NY, USA, 2019. [CrossRef]
76. Zohrevand, Z.; Glasser, U.; Shahir, H.; Tayebi, M.A.; Costanzo, R. Hidden Markov based anomaly detection for water supply systems. In Proceedings of the 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, USA, 5–8 December 2016; IEEE Computer Society: Los Alamitos, CA, USA, 2016; pp. 1551–1560. [CrossRef]
77. Ahmed, C.M.; Murguia, C.; Ruths, J. Model-Based Attack Detection Scheme for Smart Water Distribution Networks. In Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security, ASIA CCS '17, New York, NY, USA, 2–6 April 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 101–113. [CrossRef]
78. Moazeni, F.; Khazaei, J. MINLP Modeling for Detection of SCADA Cyberattacks in Water Distribution Systems. In Proceedings of the World Environmental and Water Resources Congress 2020, Henderson, NV, USA, 17–21 May 2020; pp. 340–350. [CrossRef]
79. Inoue, J.; Yamagata, Y.; Chen, Y.; Poskitt, C.M.; Sun, J. Anomaly Detection for a Water Treatment System Using Unsupervised Machine Learning. In Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW), Atlantic City, NY, USA, 14–17 November 2017; pp. 1058–1065.
80. Hindy, H.; Brosset, D.; Bayne, E.; Seeam, A.; Bellekens, X. *Improving SIEM for Critical SCADA Water Infrastructures Using Machine Learning*; Katsikas, S.K., Cuppens, F., Cuppens, N., Lambrinoudakis, C., Antón, A., Gritzalis, S., Mylopoulos, J., Kalloniatis, C., Eds.; Computer Security; Springer International Publishing: Cham, Switzerland, 2019; pp. 3–19.
81. Taormina, R.; Galelli, S. Real-Time Detection of Cyber-Physical Attacks on Water Distribution Systems Using Deep Learning. In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 469–479. [CrossRef]
82. Taormina, R.; Galelli, S. Deep-Learning Approach to the Detection and Localization of Cyber-Physical Attacks on Water Distribution Systems. *J. Water Resour. Plan. Manag.* **2018**, *144*, 04018065. [CrossRef]

83. Abokifa, A.A.; Haddad, K.; Lo, C.S.; Biswas, P. Detection of Cyber Physical Attacks on Water Distribution Systems via Principal Component Analysis and Artificial Neural Networks. In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 676–691. [CrossRef]
84. Abokifa, A.A.; Haddad, K.; Lo, C.; Biswas, P. Real-Time Identification of Cyber-Physical Attacks on Water Distribution Systems via Machine Learning Based Anomaly Detection Techniques. *J. Water Resour. Plan. Manag.* **2019**, *145*, 04018089. [CrossRef]
85. Giacomoni, M.; Gatsis, N.; Taha, A. Identification of Cyber Attacks on Water Distribution Systems by Unveiling Low-Dimensionality in the Sensory Data. In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 660–675. [CrossRef]
86. Pasha, M.F.K.; Kc, B.; Somasundaram, S.L. An Approach to Detect the Cyber-Physical Attack on Water Distribution System. In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 703–711. [CrossRef]
87. Brentan, B.M.; Campbell, E.; Lima, G.; Manzi, D.; Ayala-Cabrera, D.; Herrera, M.; Montalvo, I.; Izquierdo, J.; Luvizotto, E. On-Line Cyber Attack Detection in Water Networks through State Forecasting and Control by Pattern Recognition. In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 583–592. [CrossRef]
88. Chandy, S.E.; Rasekh, A.; Barker, Z.A.; Campbell, B.; Shafiee, M.E. Detection of Cyber-Attacks to Water Systems through Machine-Learning-Based Anomaly Detection in SCADA Data. In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 611–616. [CrossRef]
89. Housh, M.; Ohar, Z. Model Based Approach for Cyber-Physical Attacks Detection in Water Distribution Systems. In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 727–736. [CrossRef]
90. Housh, M.; Ohar, Z. Model-based approach for cyber-physical attack detection in water distribution systems. *Water Res.* **2018**, *139*, 132–143. [CrossRef] [PubMed]
91. Aghashahi, M.; Sundararajan, R.; Pourahmadi, M.; Banks, M.K. Water Distribution Systems Analysis Symposium: Battle of the Attack Detection Algorithms (BATADAL). In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 101–108. [CrossRef]
92. Quiñones-Grueiro, M.; Prieto-Moreno, A.; Verde, C.; Llanes-Santiago, O. Decision Support System for Cyber Attack Diagnosis in Smart Water Networks. *IFAC-PapersOnLine* **2019**, *51*, 329–334, Part of special issue: 2nd IFAC Conference on Cyber-Physical and Human Systems CPHS 2018, Miami, Florida. [CrossRef]
93. Ramotsoela, D.; Hancke, G.; Abu-Mahfouz, A. Attack detection in water distribution systems using machine learning. *Hum. Centric Comput. Inf. Sci.* **2019**, *9*, 13. [CrossRef]
94. Kadosh, N.; Frid, A.; Housh, M. Detecting Cyber-Physical Attacks in Water Distribution Systems: One-Class Classifier Approach. *J. Water Resour. Plan. Manag.* **2020**, *146*, 04020060. [CrossRef]
95. Bakalos, N.; Voulodimos, A.; Doulamis, N.; Doulamis, A.; Ostfeld, A.; Salomons, E.; Caubet, J.; Jimenez, V.; Li, P. Protecting Water Infrastructure From Cyber and Physical Threats: Using Multimodal Data Fusion and Adaptive Deep Learning to Monitor Critical Systems. *IEEE Signal Process. Mag.* **2019**, *36*, 36–48. [CrossRef]
96. Min, K.W.; Choi, Y.H.; Al-Shamiri, A.K.; Kim, J.H. Application of Artificial Neural Network for Cyber-Attack Detection in Water Distribution Systems as Cyber Physical Systems. In *Advances in Harmony Search, Soft Computing and Applications*; Kim, J.H., Geem, Z.W., Jung, D., Yoo, D.G.,Yadav, A., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 82–88.
97. Macas, M.; Wu, C. An Unsupervised Framework for Anomaly Detection in a Water Treatment System. In Proceedings of the 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA, 16–19 December 2019; pp. 1298–1305.
98. Zou, X.Y.; Lin, Y.L.; Xu, B.; Guo, Z.B.; Xia, S.J.; Zhang, T.Y.; Gao, N.Y. A Novel Event Detection Model for Water Distribution Systems Based on Data-Driven Estimation and Support Vector Machine Classification. *Water Resour. Manag.* **2019**, *33*, 4569–4581. [CrossRef]
99. Ghaeini, H.R.; Tippenhauer, N.O. HAMIDS: Hierarchical Monitoring Intrusion Detection System for Industrial Control Systems. In Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy, CPS-SPC 2016, Vienna, Austria, 28 October 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 103–111. [CrossRef]
100. Ostfeld, A.; Salomons, E.; Ormsbee, L.; Uber, J.G.; Bros, C.M.; Kalungi, P.; Burd, R.; Zazula-Coetzee, B.; Belrain, T.; Kang, D.; et al. Battle of the Water Calibration Networks. *J. Water Resour. Plan. Manag.* **2012**, *138*, 523–532. [CrossRef]
101. Pasha, M.F.K. Development of an Effective Hybrid Method to Detect Cyber-Physical Attack on Water Distribution Systems. In Proceedings of the World Environmental and Water Resources Congress 2018, Minneapolis, MI, USA, 3–7 June 2018; pp. 410–421. [CrossRef]
102. Aggarwal, C.C. High-Dimensional Outlier Detection: The Subspace Method. In *Outlier Analysis*; Springer New York: New York, NY, USA, 2013; pp. 135–167. [CrossRef]
103. Breunig, M.M.; Kriegel, H.P.; Ng, R.T.; Sander, J. LOF: Identifying Density-Based Local Outliers. In Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, SIGMOD '00, Dallas, TX, USA, 16–18 May 2000; Association for Computing Machinery: New York, NY, USA, 2000; pp. 93–104. [CrossRef]
104. Kang, E.; Adepu, S.; Jackson, D.; Mathur, A.P. Model-Based Security Analysis of a Water Treatment System. In Proceedings of the 2016 IEEE/ACM 2nd International Workshop on Software Engineering for Smart Cyber-Physical Systems (SEsCPS), Austin, TX, USA, 16 May 2016; pp. 22–28.

105. Patlolla, S.S.; McMillin, B.; Adepu, S.; Mathur, A. An Approach for Formal Analysis of the Security of a Water Treatment Testbed. In Proceedings of the 2018 IEEE 23rd Pacific Rim International Symposium on Dependable Computing (PRDC), Taipei, Taiwan, 4–8 December 2018; pp. 115–124.
106. Howser, G.; McMillin, B. A Modal Model of Stuxnet Attacks on Cyber-physical Systems: A Matter of Trust. In Proceedings of the 2014 Eighth International Conference on Software Security and Reliability (SERE), San Francisco, CA, USA, 30 June–2 July 2014; pp. 225–234. [CrossRef]
107. Liau, C.-J. Belief, information acquisition, and trust in multi-agent systems—A modal logic formulation. *Artif. Intell.* **2003**, *149*, 31–60. [CrossRef]
108. Mishra, V.K.; Palleti, V.R.; Mathur, A. A modeling framework for critical infrastructure and its application in detecting cyber-attacks on a water distribution system. *Int. J. Crit. Infrastruct. Prot.* **2019**, *26*, 100298. [CrossRef]
109. Douglas, H.C.; Taormina, R.; Galelli, S. Pressure-Driven Modeling of Cyber-Physical Attacks on Water Distribution Systems. *J. Water Resour. Plan. Manag.* **2019**, *145*, 06019001. [CrossRef]
110. Moraitis, G.; Nikolopoulos, D.; Bouziotas, D.; Lykou, A.; Karavokiros, G.; Makropoulos, C. Quantifying Failure for Critical Water Infrastructures under Cyber-Physical Threats. *J. Environ. Eng.* **2020**, *146*, 04020108. [CrossRef]
111. Jeong, H.S.; Abraham, D.M.; Qiao, J.; Lawley, M.A.; Richard, J.P.P.; Yih, Y. Issues in Risk Management of Water Networks Against Intentional Attacks. In Proceedings of the ASCE Pipeline Division Specialty Congress—Pipeline Engineering and Construction, San Diego, CA, USA, 1–4 August 2004; pp. 1–10. [CrossRef]
112. Shin, S.; Lee, S.; Burian, S.J.; Judi, D.R.; McPherson, T. Evaluating Resilience of Water Distribution Networks to Operational Failures from Cyber-Physical Attacks. *J. Environ. Eng.* **2020**, *146*, 04020003. [CrossRef]
113. Mathur, A. SecWater: A Multi-Layer Security Framework for Water Treatment Plants. In Proceedings of the 3rd International Workshop on Cyber-Physical Systems for Smart Water Networks, CySWATER '17, Pittsburgh, PA, USA, 21 April 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 29–32. [CrossRef]
114. Luiijf, E.; Ali, M.; Zielstra, A. Assessing and Improving SCADA Security in the Dutch Drinking Water Sector. In *Critical Information Infrastructure Security*; Setola, R., Geretshuber, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 190–199.
115. Falliere, N.; Murchu, L.O.; Chien, E. *SCADA Security Good Practices for the Drinking Water Sector*; TNO Defence, Security and Safety; Report: TNO-DV 2008 C096; TNO: Den Haag, The Netherlands, 2008.
116. Burghouwt, P.; Maris, M.; van Peski, S.; Luiijf, E.; van de Voorde, I.; Spruit, M. Cyber Targets Water Management. In *Critical Information Infrastructures Security*; Havarneanu, G., Setola, R., Nassopoulos, H., Wolthusen, S., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 38–49.
117. Adepu, S.; Mishra, G.; Mathur, A. Access Control in Water Distribution Networks: A Case Study. In Proceedings of the 2017 IEEE International Conference on Software Quality, Reliability and Security (QRS), Prague, Czech Republic, 25–29 July 2017; pp. 184–191.
118. Lenzi, S.; Terenghi, G.; Taormina, R.; Galelli, S.; Ciuccarelli, P. Disclosing cyber attacks on water distribution systems: An experimental approach to the sonification of threats and anomalous data. In Proceedings of the International Conference on Auditory Display, Tyne, UK, 23–27 June 2019.
119. Tatbul, N.; Lee, T.J.; Zdonik, S.; Alam, M.; Gottschlich, J. Precision and Recall for Time Series. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, NIPS 2018, Denver, CO, USA, 3–8 December 2018; Curran Associates Inc.: Red Hook, NY, USA, 2018; pp. 1924–1934.
120. Kurakin, A.; Goodfellow, I.; Bengio, S. Adversarial Machine Learning at Scale. *arXiv* **2016**, arXiv:1611.01236.
121. Madry, A.; Makelov, A.; Schmidt, L.; Tsipras, D.; Vladu, A. Towards Deep Learning Models Resistant to Adversarial Attacks. *arXiv* **2019**, arXiv:1706.06083.